

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 08:42:29 ; Search time 10148 Seconds
(without alignments)
10920.465 Million cell updates/sec

Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gcataaaggaaatggagc.....tcttttcaaatccgaaaaa 1733

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_env.*
- 2: gb_pat.*
- 3: gb_ph.*
- 4: gb_pl.*
- 5: gb_pr.*
- 6: gb_ro.*
- 7: gb_sts.*
- 8: gb_sy.*
- 9: gb_un.*
- 10: gb_vl.*
- 11: gb_ov.*
- 12: gb_hg.*
- 13: gb_in.*
- 14: gb_om.*
- 15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1733	100.0	1733	2	AX366075 Sequence
2	1729.8	99.8	1758	4	AF406732 Euphorbia
3	410	23.7	1538	2	CS223906 Sequence
4	400.4	23.1	1673	2	CS223962 Sequence
5	400.4	23.1	1673	2	CS223964 Sequence
6	400.4	23.1	1673	2	CS223984 Sequence
7	399.4	23.0	1781	2	E63735 Cytochrome
8	399.4	23.0	1781	2	AR205821 Sequence
9	399.4	23.0	1781	4	AF122821 Capsicum
10	398.8	23.0	1673	2	CS223966 Sequence
11	393.2	22.7	1610	2	CS223950 Sequence
12	393.2	22.7	1610	2	CS224014 Sequence
13	388.8	22.4	1566	2	CS223944 Sequence
14	388.8	22.4	1576	2	CS223902 Sequence
15	388.8	22.4	1581	2	CS223904 Sequence
16	388.8	22.4	1664	2	CS223952 Sequence
17	388.6	22.4	1657	2	CS223948 Sequence
18	387.6	22.4	1701	4	STU296346 Solanum t

19	387.4	22.4	1789	4	DQ335785	DQ335785 Medicago
20	384.8	22.2	1754	4	GM450CP3	Y10490 G. max mRNA
21	384.2	22.2	1559	2	CS224040	CS224040 Sequence
22	376.8	21.7	1509	4	AB028462	AB028462 Petunia x
23	376.6	21.7	1784	2	CS249182	CS249182 Sequence
24	369.8	21.3	1665	2	A29004	A29004 pCGP147 DNA
25	369.8	21.3	1665	2	I47318	I47318 Sequence 29
26	367.4	21.2	1666	2	AR030921	AR030921 Sequence
27	367.4	21.2	1666	2	I28313	I28313 Sequence 24
28	366.6	21.2	1800	4	GM450CP7	Y10493 G. max mRNA
29	364.6	21.0	1641	4	AF000403	AF000403 Lotus jap
30	361	20.8	1506	4	AY596975	AY596975 Centauriu
31	357.6	20.6	1572	2	CS223930	CS223930 Sequence
32	355.6	20.5	1515	4	AF368376	AF368376 Nicotiana
33	351.8	20.3	1682	4	CRO238612	AJ238612 Catharan
34	350.4	20.2	1641	4	SCU48434	U48434 Solanum cha
35	347.4	20.0	1599	2	CS223912	CS223912 Sequence
36	344.2	19.9	1593	2	CS223910	CS223910 Sequence
37	344.2	19.9	1694	2	CS224048	CS224048 Sequence
38	339.6	19.6	1823	2	CS223960	CS223960 Sequence
39	338	19.5	1767	2	CS223956	CS223956 Sequence
40	336.4	19.4	1736	2	CS223958	CS223958 Sequence
c	315	18.2	132419	4	CR954196	CR954196 Medicago
42	308	17.8	1691	4	AF022459	AF022459 Glycine m
43	296.8	17.1	1564	2	CS223914	CS223914 Sequence
44	295.2	17.0	1560	2	CS223916	CS223916 Sequence
45	294.6	17.0	1541	4	AY192573	AY192573 Catharan

ALIGNMENTS

RESULT 1
AX366075
LOCUS AX366075
DEFINITION Sequence 1 from Patent WO208269.
ACCESSION AX366075
VERSION AX366075.1 GI:18697502
KEYWORDS
SOURCE Euphorbia lagascae
ORGANISM Euphorbia lagascae
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiae; Euphorbia.

REFERENCE
AUTHORS Cahoon,E.B.
TITLE A cytochrome p450 enzyme associated with the synthesis of _g(d)712
-epoxy groups in fatty acids of plants
JOURNAL Patent: WO 0208269-A 1 31-JAN-2002;
E. I. du Pont de Nemours and Company (US)
FEATURES
source
1. .1733
/organism="Euphorbia lagascae"
/mol_type="unassigned DNA"
/db_xref="taxon:54672"

ORIGIN
Query Match 100.0%; Score 1733; DB 2; Length 1733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCATAAAGGAAATGGAGCAGAAAATCTCTTTCCGAGCATTTTAATAGTTTCT	60
Db	1	GCATAAAGGAAATGGAGCAGAAAATCTCTTTCCGAGCATTTTAATAGTTTCT	60
QY	61	GCTTGTGTTTAATCTTAGTAGTAGTCATGAGGTGTGGAAGAAACAGAAATCCACCTCCAGG	120
Db	61	GCTTGTGTTTAATCTTAGTAGTAGTCATGAGGTGTGGAAGAAACAGAAATCCACCTCCAGG	120
QY	121	GCCATGGGAAGTTTCTTATCATAGGTAATCTTCCTCATTTTATTCATCTTCTGATCTAGG	180
Db	121	GCCATGGGAAGTTTCTTATCATAGGTAATCTTCCTCATTTTATTCATCTTCTGATCTAGG	180

Qy	181	CCATGAACGTTTTAGAGCCTTTGGCTCAAAATTTATGGACCTGTTATGAGCTCTTCAAAATTGG	240
Db	181	CCATGAACGTTTTAGAGCCTTTGGCTCAAAATTTATGGACCTGTTATGAGCTCTTCAAAATTGG	240
Qy	241	CCAAGTTTCAGTGTGTGTCATTTCTTCAAGCTGAAGCAGCCAAAGAGGTTATGAAAACCTCA	300
Db	241	CCAAGTTTCAGTGTGTGTCATTTCTTCAAGCTGAAGCAGCCAAAGAGGTTATGAAAACCTCA	300
Qy	301	GGCTGATGCCCTTCGCCCAACGCCCTTATCGTCTTGAGCGCACAGATGGTTTTTATAATCG	360
Db	301	GGCTGATGCCCTTCGCCCAACGCCCTTATCGTCTTGAGCGCACAGATGGTTTTTATAATCG	360
Qy	361	GAAGATGCTCTGTTGTTTCATATGAGATCACTGGAGGAGAGTGAAGAAAATTTGGAT	420
Db	361	GAAGATGCTCTGTTGTTTCATATGAGATCACTGGAGGAGAGTGAAGAAAATTTGGAT	420
Qy	421	ACTTGAATTTCTGAGTGCACAAAAGTTCAATCCTCCAGGTTAATCCGAGAGGAGAAAT	480
Db	421	ACTTGAATTTCTGAGTGCACAAAAGTTCAATCCTCCAGGTTAATCCGAGAGGAGAAAT	480
Qy	481	GGAGATGCCATCACATTCCTCGTTTCGAAAGCCGGATCTCCGGTCAATATTTACAAAGAT	540
Db	481	GGAGATGCCATCACATTCCTCGTTTCGAAAGCCGGATCTCCGGTCAATATTTACAAAGAT	540
Qy	541	CATTATATGGCATTAATAATTTGATCATGATGAAGAACATCCGTTGGTAAATTTGTAAGCAAAA	600
Db	541	CATTATATGGCATTAATAATTTGATCATGATGAAGAACATCCGTTGGTAAATTTGTAAGCAAAA	600
Qy	601	AGAAAGATTGCTGAGTGTTCGGATGCGATCAATGAGCAGCGAGGTTTGGCACCGC	660
Db	601	AGAAAGATTGCTGAGTGTTCGGATGCGATCAATGAGCAGCGAGGTTTGGCACCGC	660
Qy	661	AGACGCTTTTCGCGCTGGAATTAATCTCACTATATCATTTGAGCTGAGTCAAAACCCAG	720
Db	661	AGACGCTTTTCGCGCTGGAATTAATCTCACTATATCATTTGAGCTGAGTCAAAACCCAG	720
Qy	721	CGGTTTGTCATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA	780
Db	721	CGGTTTGTCATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA	780
Qy	781	TAAGCCTTTTGAGCGGATACTTAATGATGTTCTATTGATCTTCAAAAATGGAAA	840
Db	781	TAAGCCTTTTGAGCGGATACTTAATGATGTTCTATTGATCTTCAAAAATGGAAA	840
Qy	841	CGTTCCAGTGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGC	900
Db	841	CGTTCCAGTGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGC	900
Qy	901	CGGAGCGCAAAACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC	960
Db	901	CGGAGCGCAAAACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC	960
Qy	961	AACTGAACCTTAGAAAAGCACAGAGAGAGTTAGACAAGTATTGTTGGTGAATGGGAAAAGT	1020
Db	961	AACTGAACCTTAGAAAAGCACAGAGAGAGTTAGACAAGTATTGTTGGTGAATGGGAAAAGT	1020
Qy	1021	TGATGAATCAAGATTTTCATGATTTTGAAATTTCTTCAAGTTAGTGGTTAAAGAACTCTAAG	1080
Db	1021	TGATGAATCAAGATTTTCATGATTTTGAAATTTCTTCAAGTTAGTGGTTAAAGAACTCTAAG	1080
Qy	1081	ATTATACCTCCGGTGTCTTGAATTCGAGGAGGTGTAGAGAAAACAAACGAAATTTGATGG	1140
Db	1081	ATTATACCTCCGGTGTCTTGAATTCGAGGAGGTGTAGAGAAAACAAACGAAATTTGATGG	1140
Qy	1141	ATATGAATTTATCCGAAACACTCGAATGTTCTGTAATCTTCGGCGATAGGAAGATCC	1200
Db	1141	ATATGAATTTATCCGAAACACTCGAATGTTCTGTAATCTTCGGCGATAGGAAGATCC	1200
Qy	1201	TAATACTTGGTTCGGAACCTTGGAAAAGTTTAAACCAGAAAAGTTTAAAGATTTGTGCAATTGA	1260
Db	1201	TAATACTTGGTTCGGAACCTTGGAAAAGTTTAAACCAGAAAAGTTTAAAGATTTGTGCAATTGA	1260
Qy	1261	TTATAAAGGGACGACATTTGAACTGGTACCATTTGGTGCAGGAAAAAGAAATATGTCCTGG	1320
Db	1261	TTATAAAGGGACGACATTTGAACTGGTACCATTTGGTGCAGGAAAAAGAAATATGTCCTGG	1320
Qy	1321	CATTACTTCAGTATTTACCAATTTGGAGTAGTGCATTATATAATCTTATTATATCATTTTAA	1380
Db	1321	CATTACTTCAGTATTTACCAATTTGGAGTAGTGCATTATATAATCTTATTATATCATTTTAA	1380
Qy	1381	TTGGAACTGGCCGATGGAATTTACACCTCAAAACCTTGATATGACTGAAGCTATTGGCGG	1440
Db	1381	TTGGAACTGGCCGATGGAATTTACACCTCAAAACCTTGATATGACTGAAGCTATTGGCGG	1440
Qy	1441	TGCTCTCAGGAAAAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGG	1500
Db	1441	TGCTCTCAGGAAAAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGG	1500
Qy	1501	CTCAAAATATTTCTTGATTACATAGAGGTTTGAATATATATAATAAACTTTTAAATAACG	1560
Db	1501	CTCAAAATATTTCTTGATTACATAGAGGTTTGAATATATATAATAAACTTTTAAATAACG	1560
Qy	1561	ATGTTCTAATATGTTTGGGTAGTTAATAGGTTTCCACCGATCATATAAGTAGGCT	1620
Db	1561	ATGTTCTAATATGTTTGGGTAGTTAATAGGTTTCCACCGATCATATAAGTAGGCT	1620
Qy	1621	TCCTTGATGATGGTTAGATTATAAGTGTGGTGGTGGATTTTATAGATGGGTTAAAT	1680
Db	1621	TCCTTGATGATGGTTAGATTATAAGTGTGGTGGTGGATTTTATAGATGGGTTAAAT	1680
Qy	1681	GATTGGATGATTAATAATAAATGAAATGTTTCTTTTCCAAATCCGAAAAA	1733
Db	1681	GATTGGATGATTAATAATAAATGAAATGTTTCTTTTCCAAATCCGAAAAA	1733
RESULT 2			
AF406732			
LOCUS	AF406732	1758 bp	mRNA linear
DEFINITION	Euphorbia lagascae cytochrome P450 (CYP726A1)		PLN 13-FEB-2002
ACCESSION	AF406732.1		
VERSION	AF406732.1		GI:18157658
KEYWORDS			
SOURCE	Euphorbia lagascae		
ORGANISM	Euphorbia lagascae		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malpighiales; Euphorbiales; Euphorbiaceae; Euphorbioideae; Euphorbiae; Euphorbia.		
REFERENCE	1 (bases 1 to 1758)		
AUTHORS	Cahoon, E.B., Ripp, K.G., Hall, S.E. and McGonigle, B.		
TITLE	Transgenic production of epoxy fatty acids by expression of a cytochrome P450 enzyme from Euphorbia lagascae seed		
JOURNAL	Plant Physiol.	128 (2),	615-624 (2002)
PUBMED	11842164		
REFERENCE	2 (bases 1 to 1758)		
AUTHORS	Cahoon, E.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-AUG-2001)		DuPont Crop Genetix, E.I. DuPont Company,
	E402/4212, Wilmington, DE 19880-0402, USA		
FEATURES	Location/Qualifiers		
source	1..1758		
	/organism="Euphorbia lagascae"		
	/mol_type="mRNA"		
gene	/db_xref="taxon:54672"		
CDS	1..1758		
	/gene="CYP726A1"		
	21..1523		
	/note="delta12-fatty acid epoxxygenase"		
	/product="cytochrome P450"		
	/codon_start=1		
	/protein_id="AAL62063.1"		
	/db_xref="GI:18157659"		
	/translation="MEQKNLSFPSPILISGLILVVMRLWKKNPPPPKPFPIIG		
	NLPULLTSLDGLHFRFALAIQIYGVMSLQIGVSAVVISSEAAAEKVNKTDAPAQ		

RPVLDIAQIVFYNRKDVLPASYGDHWRQMKKIWILFLSAKKVQSSRLIREBEMDAI
TFRLSKAGSPVNIITKIIYIGIISIMRITSVGNCKOKERILLISVADAVNEAATSGTADA
FPTMKLLHYIIIGAESKPRLHOEIIDILEIINHEKANKPFEADINLMDVLLNLOKNGN
VVPVTNESI KASVLQMF TAGSETTSKATEWMAELMKNPTFLRKAQBEVRQVFGEMG
KYDESRFDLKFPLVVKETLRLHPVVLIPRECHRETRIDGTEHPNTRIVVNAWAI
GDPNTWSEFGNFERFKDCAIDRYKGTTFELVFPFGAGKRICPGITSAITNLEJYIIN
LLYHFNWELADGITPTQDLMTEAIGALRKIDKLPIPIYQVSLGSNIS"

ORIGIN

Query Match	99.8%;	Score 1729.8;	DB 4;	Length 1758;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1731;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	GCATAAAAGGAAATGGAGCAGAAAAATCTCTCTTTCCGAGCATTTTAATAAGTTTCT	60	
DB	8	CGAGAAAAGGAAATGGAGCAGAAAAATCTCTTTTCCGAGCATTTTAATAAGTTTCT	67	
QY	61	GCTTGTTTAACTTAGTAGTAGTCAATGAGGTTGTGGAGAAACAGAAATCCACCTCCAGG	120	
DB	68	GCTTGTTTAACTTAGTAGTAGTCAATGAGGTTGTGGAGAAACAGAAATCCACCTCCAGG	127	
QY	121	GCATGGAAGTTTCTCATCATAGGTAATCTCTCTCATTTTACTCACTTCTGTGATCTAGG	180	
DB	128	GCATGGAAGTTTCTCATCATAGGTAATCTCTCTCATTTTACTCACTTCTGTGATCTAGG	187	
QY	181	CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGG	240	
DB	188	CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGG	247	
QY	241	CCAAGTTTCAGCTGTGTTCATTTCTTCACTGAAGCAGCAAGAGGTTATGAAACTCA	300	
DB	248	CCAAGTTTCAGCTGTGTTCATTTCTTCACTGAAGCAGCAAGAGGTTATGAAACTCA	307	
QY	301	GGCTCATGCTTCGCCCAAGCCCTATCGTCTTGGACGACAGATGTGTTTTATAATCG	360	
DB	308	GGCTCATGCTTCGCCCAAGCCCTATCGTCTTGGACGACAGATGTGTTTTATAATCG	367	
QY	361	GAAGATGCTTGTGTGTTTCATATGAGATCACTGGAGGCAGATGAAGAAAATTTGGAT	420	
DB	368	GAAGATGCTTGTGTGTTTCATATGAGATCACTGGAGGCAGATGAAGAAAATTTGGAT	427	
QY	421	ACTTGAATTTCTGAGTGCCAAAAAGTTCAATCTCCAGGTTAAATCCGAGAGGAAAT	480	
DB	428	ACTTGAATTTCTGAGTGCCAAAAAGTTCAATCTCCAGGTTAAATCCGAGAGGAAAT	487	
QY	481	GGAGATGCCATCAATTCCTCCGTTGNAAGCCGATCTCCGTTCAATATTTACAAAGAT	540	
DB	488	GGAGATGCCATCAATTCCTCCGTTGNAAGCCGATCTCCGTTCAATATTTACAAAGAT	547	
QY	541	CATTTATGGCATTTATAATTTTCGATCATGATAGAACATCCGTTGGTAAATTTGAAGCAAAA	600	
DB	548	CATTTATGGCATTTATAATTTTCGATCATGATAGAACATCCGTTGGTAAATTTGAAGCAAAA	607	
QY	601	AGAAAGATTGCTGAGTTGCGGATGCGAGTCAATGAGGAGCAGCAGTTTGGCACCGC	660	
DB	608	AGAAAGATTGCTGAGTTGCGGATGCGAGTCAATGAGGAGCAGCAGTTTGGCACCGC	667	
QY	661	AGACGTTTTCCGAGCTGGAAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAG	720	
DB	668	AGACGTTTTCCGAGCTGGAAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAG	727	
QY	721	CGGTTTGCATCAGGAGATTGAGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA	780	
DB	728	CGGTTTGCATCAGGAGATTGAGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA	787	
QY	781	TAAGCCTTTTGAAGCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAATGGAAA	840	
DB	788	TAAGCCTTTTGAAGCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAATGGAAA	847	
QY	841	CGTTCAGTGCCAGTGACAAACGAAAGCATCAAAGCATCCGTTTGGAAAATGTTTACTGC	900	
DB	848	CGTTCAGTGCCAGTGACAAACGAAAGCATCAAAGCATCCGTTTGGAAAATGTTTACTGC	907	

QY	901	CGGAGCGAAAAACAATCTTGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC	960	
DB	908	CGGAGCGAAAAACAATCTTGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC	967	
QY	961	AACTGAACTTAAGAAAAAGCAAGAAAGCTTGAAGAAATTTTGGTGAATGGGAAAAAGT	1020	
DB	968	AACTGAACTTAAGAAAAAGCAAGAAAGCTTGAAGAAATTTTGGTGAATGGGAAAAAGT	1027	
QY	1021	TCATGAATCAAGATTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAACTCTAAG	1080	
DB	1028	TCATGAATCAAGATTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAACTCTAAG	1087	
QY	1081	ATTATCATCTCCGGTTGTCTTCATTTCCGAGGAGGTGTAGAGAAACAAACGAATTCATGG	1140	
DB	1088	ATTATCATCTCCGGTTGTCTTCATTTCCGAGGAGGTGTAGAGAAACAAACGAATTCATGG	1147	
QY	1141	ATATGAAATTCATCCGAAACACTTCGAATTTGTGTAATGTTGGGCGATAGGAAGATCC	1200	
DB	1148	ATATGAAATTCATCCGAAACACTTCGAATTTGTGTAATGTTGGGCGATAGGAAGATCC	1207	
QY	1201	TAATACTTTGGTCGGAAACCTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGCAATTGA	1260	
DB	1208	TAATACTTTGGTCGGAAACCTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGCAATTGA	1267	
QY	1261	TTATAAAGGAGCAGACATTTGAACTGGTACCATTTGGTCAGGAAAGAAATATGTCCTGG	1320	
DB	1268	TTATAAAGGAGCAGACATTTGAACTGGTACCATTTGGTCAGGAAAGAAATATGTCCTGG	1327	
QY	1321	CATTACTTCAGCTATTACCAATTTGGAGTATGTCATTAATAATCTATTATATCATTTTAA	1380	
DB	1328	CATTACTTCAGCTATTACCAATTTGGAGTATGTCATTAATAATCTATTATATCATTTTAA	1387	
QY	1381	TTGGGAATCGGCCGATGGAATTTACACCTCAACACTTGATGACTGAAGCTATTGGCGG	1440	
DB	1388	TTGGGAATCGGCCGATGGAATTTACACCTCAACACTTGATGACTGAAGCTATTGGCGG	1447	
QY	1441	TGCTCTCAGGAAAAAATAGATCTTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGG	1500	
DB	1448	TGCTCTCAGGAAAAAATAGATCTTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGG	1507	
QY	1501	CTCAAAATTTCTTGATTCATAGAGGTTGAAATATATATATAATAATAAATTTAAATTAACG	1560	
DB	1508	CTCAAAATTTCTTGATTCATAGAGGTTGAAATATATATAATAATAAATTTAAATTAACG	1567	
QY	1561	ATGTTCTTAATATGTTGGTGAGTTAATAAGGTTTCCACCGATCATATAAGTAGCCT	1620	
DB	1568	ATGTTCTTAATATGTTGGTGAGTTAATAAGGTTTCCACCGATCATATAAGTAGCCT	1627	
QY	1621	TCTTTGATGGATGGTTAGATTATTAATCAGTTGTGGTTGGATTTTGTAGATGGTTAAAT	1680	
DB	1628	TCTTTGATGGATGGTTAGATTATTAATCAGTTGTGGTTGGATTTTGTAGATGGTTAAAT	1687	
QY	1681	GATTTGGATGATATAATAATAATGAAATGTTTCTTTTCAATCCGAAAAA 1733		
DB	1688	GATTTGGATGATATAATAATAATGAAATGTTTCTTTTCAATCCGAAAAA 1740		

RESULT 3
CS223906

LOCUS CS223906 1538 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 162 from Patent WO2005111217.

ACCESSION CS223906
VERSION CS223906.1 GI:83685434

KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum

REFERENCE
AUTHORS Xu, D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 162 24-NOV-2005;

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Solanales; Solanaceae; Nicotiana.

FEATURES		U.S. Smokeless Tobacco Company (US)	
source		Location/Qualifiers	
		1. 1538	
		/organism="Nicotiana tabacum"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:4097"	
ORIGIN			
Query Match		23.1%; Score 410; DB 2; Length 1538;	
Best Local Similarity		58.1%; Pred. No. 5e-116;	
Matches		816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;	
Qy	113	CCTCAGGGCCATGGAAGTTTCCTATCATAGTAACTCTTCCTCATCTTATTTACTCACITCT	172
Db	129	CCACCAGGTCCATGGAAACTACCAATACTAGGAAGTATGCTTCATATGGT-----TGGT	182
Qy	173	GATCTAGGCCATGAACGTTTGTAGAGCTTGGCTCAAAATTTATGGACCTGTTTATGAGTCTT	232
Db	183	GGACTACACACCATGTCTTAGAGATTATGGCAAAAATATGACCGCTTATGCACCTT	242
Qy	233	CAAAATGGCCAAAGTTTCAGCTGTGTCTCAATTTCTTCAGCTGAAGCAGCAAGAGGTTATG	292
Db	243	CAAATAGGTGAAGTTTTCGAGTTTGTGGTTACTTCTCTGATATGGCAAAAAGAGTACTA	302
Qy	293	AAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGAGCGACAGATTGTGTTT	352
Db	303	AAAACCTCATGATCGCTTTCGCGTCTAGGCGCTAGCCCTTTTGGCCCCGGAGATTGTCTGT	362
Qy	353	TATAATCGGAAGATGTCTGTTGCTTCATATGAGATCACTGGAGCGACAGTGAAGAAA	412
Db	363	TACAATAGTCTGATCTTCGCTTTTGGCCCTATGCGGATTTATGGAGCAAAATGCGTAAA	422
Qy	413	ATTTGGATACITTTGATGTCGCAAAAGTTCAATCCTCAGGTTAATCCGAGAG	472
Db	423	ATATGTGCTTGGAAAGTGTCAAGTGCCAAAGATGTTGCGACATATAGCTCTATTAGGCGC	482
Qy	473	GNAGAAATGGAGGATGCCATCACATTCCTCGTTGGAAGCCGATCTCCGGTCAATATT	532
Db	483	GATGAAGTTCTTCGCTCCTTAATTTTATCCGGTCACTCTTCGTGGAGCCTGTTAATATT	542
Qy	533	ACAAAGATCATTTATGGCAATTAATTTGATCATGATAAGAAACATCCGTTGGTAA---T	589
Db	543	ACGGAAGATCTTTTGTTCACAAGTCCATGACATGTAGATCAGCGTTTGGGCAAGTA	602
Qy	590	TGTAGCAAAAAGAAAGATTGCTGAGTGTGCGGATGCAAGTCAATGAGCGACGAGT	649
Db	603	TTCAAGGAGCAAGCAAAATTTATACAACTAATTAAGAAAGTTATACTCTTAGCAGAGGG	662
Qy	650	TTTGGCAGCGACAGCGCTTTCCGAGTGGAAATTTACTTCACTATATCATTTGGAGCTGAG	709
Db	663	TTTGAATGTGGCTGACATATTCCTTCATACAAAGTCTCTTCATGTGCTCAGTGGAAATGAG	722
Qy	710	TCAAAACCCAGGCGTTTGCATCAGAGATTGACGATATACCTTTGAAGAGATTCTTAAATGAA	769
Db	723	GGTAGATTATGAATGCACACCATTAAGGTAGATGCTATTGTTGAGNATGTCATCAACGAG	782
Qy	770	CAC-----AAAGCCAATAAGCCCTTTTGAAGCGGATACTTAATG	808
Db	783	CACAAGAAAAATCTTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT	842
Qy	809	GATGTTCTATTGAATCTTCAAAAATGGAACGTTCCAGTCCGAGTCGAGTGAACCAAGC	868
Db	843	GATGTTCTTCTAAAATTTATGAATGATGGAGGCCCTTCAATTTCTTATCAACCAACGACAC	902
Qy	869	ATCAAGACATCCGTTTTCGAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA	928
Db	903	ATCAAGCTATATCTTTGACATGTTTGTGCTGGAACAGAGACTTCATCGTCAACAATT	962
Qy	929	GAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAAGTAAGAAAAAGCAACAAGAA	988
Db	963	GTGTGGGCTATGGTGGAAATGGTGAATAATCCAACTGTATTGCGAAAGCTCAAGCAGAA	1022
Qy	989	GTTAGCAAGTATTTGGTGAAATGGGAAAGTTGATGAATCAAGATTTTCATGATTGAAA	1048
Db	1023	GTAAGAGATGCATTTTAGAAAAAAGAAACTTTTGTGAAATGATGTGGAGAGCTAAAC	1082
Qy	1049	TTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTACATCTCCGGTT---GTCTTGATT	1105
Db	1083	TATCTAAAGTTTAGTCAITTAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC	1142
Qy	1106	CCGAGGAGTGTAGAGAAAACAACAGCAATTGATGGATATGAATTCATCCGACACTCGA	1165
Db	1143	CCAAGAGAATGTAGGGAAGAGACAAATATAAACCGCTACACTATTCCTGTAAAGACCAAA	1202
Qy	1166	ATTCTTGTGAATGCTTGGGCGATAGGAAGAGATCCTAATATCTTGTCTCGGAACCTGGAAG	1225
Db	1203	GTCAFGTGAATGTTTGGGCATTCGGAAGAGATCCAAAATATTGGGATGATGCAGAACT	1262
Qy	1226	TTTAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTATAAAGGACGACATTTGAACGT	1285
Db	1263	TTTAAGCCAGAGAGATTGAGCAGTGTCTAAGGATTTTGTGTGTAATAATTTTGAATAT	1322
Qy	1286	GTACCAATTTGGTCAGGAAAGAAATATGTCCTCGCATTTACTTCAGCTATTACCAATTTG	1345
Db	1323	CTTCCATTTGGTGTGGAAGGAGATTGTCCAGGATTTGTTTGGTTTACCTAATGCT	1382
Qy	1346	GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGGAACTGGCCGATGGAATTACA	1405
Db	1383	TATTTGCCATTTGGCTCAATTTACTTTATCACCTTTGATTGGGAACCTCCCACTGGAATCAA	1442
Qy	1406	CCTCAAAACATTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT	1465
Db	1443	CCAAAGCACTTTGACTTTGACTGAGTTGGTTGGAGTAACTGCCGCTAGAAAAAAGTGACCTT	1502
Qy	1466	AAHTTGATTCCTATTCCATATCAA	1489
Db	1503	TACTTGGTTGGACTCTCTTATCAA	1526
RESULT 4			
CS223962		1673 bp DNA linear	PAT 15-DEC-2005
LOCUS		Sequence 218 from Patent WO2005111217.	
DEFINITION		CS223962	
ACCESSION		CS223962	
VERSION		CS223962.1	GI:83685860
KEYWORDS		Nicotiana tabacum (common tobacco)	
SOURCE		Nicotiana tabacum	
ORGANISM		Nicotiana tabacum	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.	
AUTHORS		Xu,D.	
TITLE		Nicotiana nucleic acid molecules and uses thereof	
JOURNAL		Patent: WO 200511217-A 218 24-NOV-2005;	
FEATURES		U.S. Smokeless Tobacco Company (US)	
source		Location/Qualifiers	
		1. 1673	
		/organism="Nicotiana tabacum"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:4097"	
ORIGIN			
Query Match		23.1%; Score 400.4; DB 2; Length 1673;	
Best Local Similarity		57.7%; Pred. No. 5e-113;	
Matches		810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;	
Qy	113	CCTCAGGGCCATGGAAGTTTCCTATCATAGTAACTCTTCCTCATCTTATTTACTCACITCT	172
Db	135	CCACCAGGTCCATGGAAACTACCAATACTAGGAAGTATGCTTCATATGGT-----TGGT	188
Qy	173	GATCTAGGCCATGAACGTTTGTAGAGCTTGGCTCAAAATTTATGGACCTGCTTATGAGCTTT	232
Db	189	GGACTACCAACCATGTCTCTTAGAGATTAGCCAAAATAATAGGACCACTTATGCACCTT	248
Qy	233	CAAAATGGCCAAAGTTTTCAGCTGTGTCTCAATTTCTTCAGCTGAAGCAGCAAGAGGTTATG	292

Db 249 CAATTAGGTGAAGTTTCTCGGTTGGTGTACTTCTCTGATACGGCAAAAGAAATATTA 308
Qy 293 AAAAATCAGGCTGATGCGCTTTCGCCAACGCCCTATCGTCTTGGACGACAGATTTGGTTT 352
Db 309 AAAATCATGACATCGCTTTTGGGTCTAGGCTTAGCCTTTTGGCCCGGAGATTTGCTGT 368
Qy 353 TATAATCGGAAGAAGATGTCTTGTGTCTCATATGGAGATCACTGGAGGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATAGCTTTTGGCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAGTTCAATCCTCGAGTTTAATCCGAGAG 472
Db 429 ATATGTGCTTGGAAAGTGTCTAGTGCACAAAGATTTTCGGACATTTAGCTCTATTAGGCG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATTTCTCCGTTTGGAAAGCGGATCTCCGGTCAATATT 532
Db 489 ATGAAGTCTTTCGCTCATTAATTTTATCCGGTCACTCTTCTGGTGAACCTAATTAATGTT 548
Qy 533 ACAAGATCAATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAGCAAAAGAAAGATGCTGAGTGTTCGGATGCGATGCAATGAGGACGACGAGT 649
Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGTATCTCTTAGCAGAGGG 668
Qy 650 TTTGSCCGCAGAGCGCTTTTCGACGTGGAAATTAATTTCACTATATCATTTGAGCTGAG 709
Db 669 TTTGATGTGGCTGACATATTCCTTCTCAGTAAGTTTCTTTCATGTCTCAGTGGAAATGAG 728
Qy 710 TCAAAACCCAGCGCTTTGATCAGGAGATTTGACATATATCTTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCATTAAGGTAGATGCCATTTGTTGAGNATGTCATCAATGAG 788
Qy 770 CAC-----AAAGCAATAAGCGCTTTTGAAGCGGATACTTAATG 808
Db 789 CACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGAAACGTTTCCAGTCCAGTGACAAACCAAGC 868
Db 849 GATGTTCTTGAAGACTTATGAATGATGGAGCGCTTCAATTTCTTATCACCACACACAC 908
Qy 869 ATCAAGACTATAATTTTGACATGTTTGTGTCGGGACAGAGACTTCATCGTCAACAAT 968
Db 929 GAATGGGTAAATGGCAGAGCTGATGAATAATCCAACTGAACCTAAGAAAAGCAACAAGAA 988
Qy 969 GTGTGGCTATGTTAGAAATGTTGAAAAATCCAGCGCTATTTCGCAAAAGCTCAAGCAGAA 1028
Db 989 GTTAGCAAGTATTGTTGAAATGGAAAGTTGATGATCAAGATTTTCATGATTTGAAA 1048
Qy 1029 GTAAGAGAGCAATTTAGAGGAAAAAGAACTTTTCGATGAAATGATGTGGAGGAGCTAAC 1088
Db 1049 TTTCTCAAGTTAGTGGTTAAAGAAACTCTAAGATTACATCTCCCGTT---GTCTTGATTT 1105
Qy 1089 TACCTAAGTTAGTAATAAAGAAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1148
Db 1106 CCGAGGAGGTAGAGAAACAAACGAATTTGATGATGAAATTCATCCGAACACTCGA 1165
Qy 1149 CCAAGAGAATGTAGGGAAGAGACAAATATAAAGCGCTACACTATTCTCTGTAAGACCAA 1208
Db 1166 ATTTGTTGAAATGCTTGGCGATAGAGAGATCTTAATCTTGTCTCGGACCTGGAAG 1225
Qy 1209 GTCTAGTTAATGTTTGGGCTTTGGGAAGAGATCCAAAATATTTGAAATGACGCAANAAT 1268
Db 1226 TTTAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGAGACGACATTTGAACTG 1285
Qy 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTTAGGATTTTGTGTTAATAATTTGAAATAT 1328
Db 1286 GTACCATTTGTGAGGAAAAAGAAATATGTCTCTGGCATTTACTTTCAGCTATTACCAATTTG 1345

Db 1329 CTTCCATTTGGTGGCGAAGGAGGATTTGTCTCGGATTTCTGTTGGCTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACTGGCCGATGGAATTACA 1405
Db 1389 TATTTGGCAATTTGGCTCAATTTACTATATCACTTTTCGATTTGAAACTCCCTCTGGAATCGAA 1448
Qy 1406 CCTAAACACATTTGATATGACTGAAGCTATTTCGGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCGACTTGGACTTGAATGAGTTGTTGGAGTAATTCGCCCTAGAAAAAGTACCTT 1508
Qy 1466 AAGTTGATTTCTTATTCATATCAAA 1489
Db 1509 TACTTTGGTTCGACTCTTATCAA 1532

RESULT 5
CS223964 1673 bp DNA linear PAT 15-DEC-2005
LOCUS Sequence 220 from Patent WO2005111217.
DEFINITION CS223964
ACCESSION CS223964
VERSION CS223964.1 GI:83685861
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1
AUTHORS Xu,D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 220 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
Location/Qualifiers
1..1673
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN
Query Match 23.1%; Score 400.4; DB 2; Length 1673;
Best Local Similarity 57.7%; Pred. No. 56-113;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113. CTTCCAGGCGCATGGAAGTTTCTCTATCATAGTAATCTTCTCTCATTTTACTCACTCT 172
Db 135 CCACAGCTCCATGGAACACTACCAATCTAGGAAGTATGCTTCATATGCT-----TGT 188
Qy 173 GATCTAGGCCATGAACGTTTGTAGAGCTTGGCTCAAAATTTATGACCTGTTATGAGTCTT 232
Db 189 GGACTTACCACACCATGTCTCTTAGAGATTTAGCCAAAAAATATGGACCACTTATGCACCTT 248
Qy 233 CAATTTGGCCAAAGTTTTCAGCTGTTGTCTATTTCTTTCAGCTGAAGCAGCCAAAGGTTATG 292
Db 249 CAATTTAGGTGAAGTTTCTGCGGTTGTGTTTCTCTCTGATACGCAAAAGAGATTTA 308
Qy 293 AAACTCAGGCTGATGCTCTCGCCCAACGCTATCTGCTTTCGAGCGCACAGATTTGTGTT 352
Db 309 AAACTCATGATCGCTTTTGGCTTAGGCTAGCCTTTTGGCCCGGAGATTTGCTGT 368
Qy 353 TATAATCGGAAGAAGTGTCTTGTGTTCTCATATGAGATCACTGAGGCGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATCTAGCTTTCGCGCTTATGGCGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAAAGTTTCAATCTTCAGGTTAATCCGAGAG 472
Db 429 ATATGTGCTTTGGAAAGTGTCTCAGTSCCAAGAAATGTTTCGACATTTAGCTCTATTAGGCG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATTTCTCGTTTCGAAAGCGGATCTCCCGTCAATATT 532
Db 489 ATGAGTCTTCTGCTCATTAATTTTATCCGGTCACTCTTCTGTTGAGAACCTATTAAATGTT 548
Qy 533 ACAAGATCATTTTATGGCAATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T 589

Db 549 ACGAAAGGATCTTTTGTTCACAAGCTCCATGATAGATCAGCGTTTGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTGCCGATGCAATGAGCGACACGAGT 649
Db 609 TTCAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGATCTTTAGCAGGAGG 668
Qy 650 TTTGGCCCGCAGACGCTTTTCCGACGTGGAAATTTACTTCACTATATCAATTTGGAGCTGAG 709
Db 669 TTTGATGGCTGACATATCCCTTCACTGAAGTTTCTTCACTGTCTCAGTGGAAATGAG 728
Qy 710 TCAAAACCCAGCGCTTTCATCAGGAGATTGACGATATACCTTGAAGATTTCTTAATGAA 769
Db 729 GGTAAAGATTGAATGCAACCATAGGTAGTCCATTGTGTGAAGATGTCAATCAATGAG 788
Qy 770 CAC-----AAAGCCCAATAGCCCTTTTGAAGCGGATAAATCTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACAGTTCCAGTGCCAGTGACAAACGAAGC 868
Db 849 GATGTTCTTAAGACTTATGAATGATGGAGCGCTTCAATTTCCCTATCACCACGACAAAC 908
Qy 869 ATCAAGCATCCGTTTTCGAATGTTTACTCCCGGACGGAACAACTTCCGAAAGCTACA 928
Db 909 ATCAAGCCATAATTTTGGCATGTTTCTGCGGGACAGAGCTTCATCTCAACAAT 968
Qy 929 GAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAAGAAAGCAAGAA 988
Db 969 GTGTGGGCTATGTTAGAAATGGTGAATAATCCAGCCGTTATTCGCGAAAGCTCAAGCAA 1028
Qy 989 GTTAGCAAGATTTTGGTGAATGGGAAAGTTGTAGTAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTAAGAGAGCATTTAGAGGAAAGAAACTTTTCGATGAAATATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TTCTTCAAGTTAGTGGTTTAAAGAACTCTAGATTTACATCTCCGCTT- --GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAGAGCTTCATCACCGGTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGGTAGAGAAACAAACGAATTTGATGGATATGAATTCATCCGAAACACTCGA 1165
Db 1149 CCAAGAGATGTAGGGAAGAGACAAATATAACGGCTACACTATTTCTGTAAAGACCAA 1208
Qy 1166 ATTTGTTGAATGCTTTGGCGGATAGAGAGATCCTAATATCTTGTGCGGAACCTGGAAG 1225
Db 1209 GTCATGTTAATGTTTGGGCTTTGGGAAGAGATCCAAAAATATGGAATGACGACAGAACT 1268
Qy 1226 TTTAACCCAGAAAGCTTTAAAGATTGTGCAATTTGATTATAAGGGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTTAAGGATTTTGTGGTAATTAATTTTGAATAT 1328
Qy 1286 GTACCATTGTTGGTGCAGAAAAAGAAATATGTCCTGGCATTTACTTTCAGCTATTACCAATTTG 1345
Db 1329 CTTCCATTTGTTGGTGGGGAAGGAGGATTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGATTAATAAATCTATTATATCAATTTTAATTTGGAACTGGCGATGGAATTAACA 1405
Db 1389 TATTTGCCATTGGCTCAATTAATATATATCACTTCGATTTGGAATCTCCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAACACTTGTATGATGCTGAAGCTATTGCGGTGCTCTCAGGAAAAAATAATAGATCTT 1465
Db 1449 CCAAGGACTTGGACTTGACTGAGTTGGTTGGATTAATCTGCCGCTAGAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTCTTATTCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532

RESULT 6
CS223984
LOCUS CS223984 1673 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 240 from Patent WO200511217.
ACCESSION CS223984
VERSION CS223984.1 GI:83685871

KEYWORDS
SOURCE
ORGANISM

Nicotiana tabacum (common tobacco)
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE
AUTHORS
TITLE
JOURNAL

Xu, D.
Nicotiana nucleic acid molecules and uses thereof
Patent: WO 200511217-A 240 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)

FEATURES
source

1. .1673
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN

Query Match 23.1%; Score 400.4; DB 2; Length 1673;
Best Local Similarity 57.7%; Pred. No. 5e-113;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113 CCTCAGGGCCATGGAAGTTTCCTATCATAGGTAACTTCCTCATTTATTACTCATCTCT 172
Db 135 CCACAGGTCATGGAACCTACCAATAGTGAAGTATGCTTCATATGTT-----TGCT 188
Qy 173 GATCAGGCCATGAACGTTTGTAGAGCTTGGCTCAAAATTTATGACCTGTTATGAGTCTT 232
Db 189 GGACTACCAACCATGCTCTTAGAGATTTAGCAAAAAAATATGGACCACTTATGCACCTT 248
Qy 233 CAAATTGGCCAAAGTTTCAGCTGTTGTCAATTTCTCAGCTGAAGCAGCAAGAGAGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTGCGGTTGTGGTTACTTCTCTGATACGGCAAAAGAGATTATA 308
Qy 293 AAAACTCAGGCTGATGCCCTTCGCCCAACGCCCTATCGTCTTGGAGCGCAGATTTGTGTTT 352
Db 309 AAACTCATGATCGCTTTTGGCTCAGGCTAGCCCTTTGGCCCCGAGATTTGCTGT 368
Qy 353 TATAATCGAAAGATGCTTCTGTTGCTTCATATGGAGATCACTGGAGCGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATCTAGCCCTTTGCCCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGTGTCTTGGAGTGCTCAGTGCCAAAGATGTTTCGACATTTAGCTCTATTAGGCGG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCCGTTTCGAAAGCGGATCTCCGTCATATTT 532
Db 489 AATGAAGTTCTCGTCTCATTAATTTTATCCGGTCACTCTCTGTTGAACCTATTAAATGTT 548
Qy 533 ACAAGATCATTTATGGCATTAATAATTTGATCATGATAAGNACATCCGTTGGTAA--T 589
Db 549 ACGAAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTGCGGATGCAATGAGGACGACGAGT 649
Db 609 TTCAAGAGCAAGACAAATTTATCAACTAATTAAGAAAGTGATACTCTTAGCAGAGGG 668
Qy 650 TTTGSCACCGCAGAGCGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGAGCTGAG 709
Db 669 TTTGATGTGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGGAAATGAAG 728
Qy 710 TCAAAACCCAGCGCTTTTCGATCAGGAGATTGACGATATACCTTGAAGAGATTTCTTAATGA 769
Db 729 GGTAAGATTATGAATGCACCATTAAGGTAGTGCATTTGTTGAGATGTCTCATATGAG 788
Qy 770 CAC-----AAAGCCCAATAGCCCTTTTGAAGCGGATAAATCTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAATGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACAGTTCCAGTGCCAGTGACAAACGAAGC 868
Db 849 GATGTTCTTAAGACTTATGAATGATGGAGCGCTTCAATTTCCCTATCACCACGACAAAC 908

```
QY 869 ATCAAGACATCCGTTTTCGAAATGTTTACTGCGGAGCGAACAACATCTTCGAAAGCTACA 928
    |||||
Db 909 ATCAAGACCATATATTTTGACATGTTTGTGCGGAGCAGAGACTTCATCGTCAACAATT 968
    |||||
QY 929 GAATGGTAAATGCGAGAGCTGATGAAATAATCCAACTGAACCTGAAGAAACACAAGAGAA 988
    |||||
Db 969 GTGTGGCTATGATAGAAATGGTGAATAATCCAGCCGTATTCGCGAAAGCTCAAGCAGAA 1028
    |||||
QY 989 GTTAGACAAGTATTTGGTCAATGGGAAGTGTGATGATCAAGATTTTCATGATTTGAAA 1048
    |||||
Db 1029 GTAAGAGAAGCATTTAGAGGAAAGAAACTCTTCGATGAAATGATGTGGAGGAGCTAAAC 1088
    |||||
QY 1049 TTCTTCAAGTTAGTGTGTTAAAGAACTCTAAGATTACATCTCCCGTT---GTCTTGATT 1105
    |||||
Db 1089 TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
    |||||
QY 1106 CCGAGGAGGTAGAGAAACAACACGAATGTGATGATGAAATTCATCCGAACACTCGA 1165
    |||||
Db 1149 CCAAGAGAAATGTAGGAGAGAGACAAATATAAACGGCTACACTATTCTGTAAGACCAA 1208
    |||||
QY 1166 ATGTGTGCAATGCTTTGGSCGATAGGAAGATCCCTAATCTTGTTCGGAACCTGGAAG 1225
    |||||
Db 1209 GTCATGTTAATGTTTGGGCTTTGGGAAGAGATCCAAATATGTGAATGACGAGAACT 1268
    |||||
QY 1226 TTTAACCCAGAAAGGTTTAAAGATTGTGCAATTGATTTATAAAGGAGCAGCATTTGAACTG 1285
    |||||
Db 1269 TTTATGCCAGAGAGATTGAGCAGTCTCTAAGGATTTTGTGTAATAATTTGNAATAT 1328
    |||||
QY 1286 GTACATTTGGTGCAGGAAAGAAATATGTCCTCGCATTAATTCAGCTATTACCAATTG 1345
    |||||
Db 1329 CTTCCATTTGGTGGGAGGAGGATTTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT 1388
    |||||
QY 1346 GAGTATGCAATTAATACTATTATATCATTTTAAATGGAATCTGCGGATCTCAGGAAAAATAGATCTT 1405
    |||||
Db 1389 TATTTGCCATTTGGCTCAATTAATATATCATTTTAAATGGAATCTGCGGATCTCAGGAAAAATAGATCTT 1448
    |||||
QY 1406 CTTCAAAACATTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAATAGATCTT 1465
    |||||
Db 1449 CCAAGCGACTGGACTTGACTGAGTTGTTGGAGTAACCTGCGGTAGAAAAAGTACCTT 1508
    |||||
QY 1466 AAGTTGATTCCTATTCATATCAAA 1489
    |||||
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532
    |||||

RESULT 7
E63735 LOCUS 1781 bp DNA linear PAT 27-AUG-2002
DEFINITION Cytochrome P450 gene highly expressed in incompatible interaction.
ACCESSION E63735
VERSION E63735.1 GI:22557608
KEYWORDS JP 2001231560-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE unclassified sequences.
AUTHORS 1 (bases 1 to 1781)
Jun, O.B., Kyungu, K. and Soon, K.Y.
TITLE Cytochrome P450 gene highly expressed in incompatible interaction
JOURNAL Patent: JP 2001231560-A 1 28-AUG-2001;
KOREA KUMHO PETROCHEMICAL CO LTD
COMMENT OS Pepper (Capsicum annuum)
PN JP 2001231560-A/1
PD 28-AUG-2001
PF 18-FEB-2000 JP 2000040456
PI O BONGU JUN, KOMUN KYUNGU, KIM YOUNG SOON
PC C12N15/09//A01H5/00, C07K14/415, C12N5/10, C12N15/00, C12N5/00 CC
Strandedness: Single;
CC Topology: Unknown;
CC single strand
CC topology, unknown
CC molecular type, cDNA to mRNA
FH Key Location/Qualifiers.
```

```
FEATURES
    source
    Location/Qualifiers
    1..1781
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"

ORIGIN
    Query Match 23.0%; Score 399.4; DB 2; Length 1781;
    Best Local Similarity 56.3%; Pred. No. 1e-112;
    Matches 826; Conservative 0; Mismatches 611; Indels 30; Gaps 3;

QY 99 AGAAACAGAATCCACCTCCAGGGCCATGGAACTTTCTCATCATAGTAATCTTCTCCTCAAT 158
    |||||
Db 86 AAAAGCTAAACTTACCTCTCGTCCATGGAAATACCTTTTATTGAAGCCATACATCACT 145
    |||||
QY 159 TATTACTCATCTCTGATCTAGGCCATGAACGTTTTAGAGCCCTTGGCTCAAAATTTATGGAC 218
    |||||
Db 146 TGGCAGTGGCAGGTCCTCTCATCATGCGCTAAATAATCTAGCAAAACTTTATGGGC 205
    |||||
QY 219 CTGTTTATGAGTCTTCAAAATTTGCCAAAGTTTTCAGCTGTTGTCTTCTTTCAGCTGAAGCAG 278
    |||||
Db 206 CGCTCATGCACTTACGACTCGGGGAAATTCCTACCGTCAATTTTCTCCCGCGAATGG 265
    |||||
QY 279 CCAAGAGGTTTATGAAAACTCAGGCTGATCCCTTCGCCCAACGCCCTATCGTCTTGGACG 338
    |||||
Db 266 CGAAGGAAGTACTAAAACTCAGCACTCGCTTTTCGCAACGAGGCGGAAACTTGTGTG 325
    |||||
QY 339 CACGATTTGTGTTTATTAATCGGAAAGATGCTGTTGCTTTCATATGGAGATCACTGGA 398
    |||||
Db 326 CTGACATCGTCCATTTATGATAGTACGGATATAGCAATTTTCTCATATGGTGAATACCTGGA 385
    |||||
QY 399 GGCAGATGAAGAAAAATTTGGATCTTGAATTTCTGAGTGCCAAAAAAGTTTCAATCTCTCCA 458
    |||||
Db 386 GGCAGATTCGTAAAAATTTGCACTCTCGAACTCTTGTAGGCCAAGATGGTCAAAATTTCTTTA 445
    |||||
QY 459 GGTTAATCCGAGAGAGAAATGGAAGATGCCATCACATTCCTCGTTTCGAAAGCGGAT 518
    |||||
Db 446 GCTCAATTCGCCAGGATGAGCTGTCGATGATGTCATCTATATCGAACCATGCCAAAT 505
    |||||
QY 519 CTCGGTCAATATTTACAAAGATCATTTATGGCATTTAATTTTCGATCATGATAGAAACAT 578
    |||||
Db 506 TTCCCGTCAACCTTTACAGACAAAAATTTTGGTTTTACAAGTTTCGGTAACTTGTAGATCAG 565
    |||||
QY 579 CCGTTGGTAA--TTGTAAGCAAAAGAAAGATTCGTGAGTGTTCGCCGATGCAGTCAATG 635
    |||||
Db 566 CTCGGGAAAAATATGCTGACCAAGACAACTGATTAATTTTCATGAGGGAATAATAT 625
    |||||
QY 636 AGGCAGCAGCAGATTTTGGCACCGCAGACGCTTTTTCGACGTGGAAATTTACTTCACTATA 695
    |||||
Db 626 CATTGACAGGTGGATTATGATTTGCTGATTTTTCCTTACATGGAATAATGCTTACATGATG 685
    |||||
QY 696 TCATTGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGAGATTTGACATATACTTTGAAG 755
    |||||
Db 686 TTGGTGGTTCAAAAACCTAGACTGCTGAAGGCTCATCGTAAAAATCGATGAGATTTTGGAA 745
    |||||
QY 756 AGATTCTTTAATGAACACAAAGCCAATAAGCCT-----TTTG 791
    |||||
Db 746 ATGTAGTGAATGAGCACAACAAAGATCGAGCGGATGCGCAAAAGGGTAAATGCGGAATTTG 805
    |||||
QY 792 AAGCGGATAACTTAATGGATGTTCTTATTTGAATCTTCAAAAAAATGGAAACGTTTCCAGTGC 851
    |||||
Db 806 GCGGTGAAGATTTTGATCGATGTTTGTCTAAGGGTTTCGAGAAAGTGGAGAGTTTCAAAAT 865
    |||||
QY 852 CAGTGACAAACGAAGCATCAAGCATCCGTTTTCGCAATGTTTACTGCCGGGAGCGAAA 911
    |||||
Db 866 CCATCAGCGATGACAAATATCAAAATCAATATTTAGTGGACATGTTTCCCGCTGGAATCGAAA 925
    |||||
QY 912 CAATTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACATAA 971
    |||||
Db 926 CGTATCGACAACTATTAATTTGGGCATTTAGCTGAATGATGAGAAACCAAGTGTTCCTAG 985
    |||||
QY 972 GAAAAGCACAAGAAAGTTAGACAAGTATTTGGTGAATGTTGGTGAATGCGGAAAGTGTGATGAATCAA 1031
    |||||
```

Db	986	CAAGGCACAAGCTGAAGTGAGACAAGTCTTTGAAGGAAAGAAAGGTTTTTCACAAATTTG	1045
Qy	1032	GATTTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTACATCCTC	1091
Db	1046	ATCTTGATGAGTTGAGTACTTTGAAGTTAGTAATCAAGAAACTCTAAGGATGCACCTC	1105
Qy	1092	CGGT- --TGCTTTGATTTCCGAGGAGTGTAGAGAAACAACAGAAATTTGATGGATATGAAA	1148
Db	1106	CAATTCCTCTATTAGTCCCTAGAGAAATGTATGAAGGATACAAAGATTTGACGGGTACAATA	1165
Qy	1149	TTCATCCGAACACTCGAATTTGTTGTGAATGCTTTGGCGCATAGGAGAGATCCTTAATACTT	1208
Db	1166	TACCTTTCAAACTCGAGTCATAGTTAATGCATGGGCAATTTGGACGAGATCCTGAAAAGTT	1225
Qy	1209	GGTCCGAACCTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAG	1268
Db	1226	GGGATGACCTGAAAGCTTTTCCCGACAGAGATTCGAGAAATAGTTCTGTGTGACTTTCTTG	1285
Qy	1269	GGACGACATTTGAACCTGGTACCAATTTGGTGCAGGAAAAAGAAATATGTCTGGCATTACTT	1328
Db	1286	GAAGCCATCATCAATTTATTTCATTTTGGTGGCGGAAGAGATTTGTCTTGGAAATGCTTT	1345
Qy	1329	CAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAATTGGGAAC	1388
Db	1346	TTGGTTTAGCCAATTTTGGACAACCATTAGCTCAATTTACTTTATCACCTTCGATCGGAAC	1405
Qy	1389	TGGCCGATGGAATTAACCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCA	1448
Db	1406	TCCCTAATGGACAAAGTCAGAAATTTGGACATGACGGAGTCACCTGGAAATTTCTGCA	1465
Qy	1449	GGAAAAAATAGATCTTAAAGTTGATTCCTATTTCATATCAAGTTAGCTTAGGCTCAAATA	1508
Db	1466	CAAGAAAGGATGATCTTGTTTGATTGCCACCCCTTATGATCCTTGAATGTATTGAGACA	1525
Qy	1509	TTTCTTGATTACATAGGAGGTTGAAA	1535
Db	1526	GTTGTAGAAATAAAAAAGAGGGAGAA	1552
RESULT 8			
AR205821			
LOCUS	1781 bp DNA linear PAT 20-JUN-2002		
DEFINITION	Sequence 1 from patent US 6369212.		
ACCESSION	AR205821		
VERSION	AR205821.1 GI:21503501		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1781)		
AUTHORS	Oh,B.-J., Ko,M.Kyung, and Kim,Y.Soon.		
TITLE	Cytochrome P450 gene highly expressed in the incompatible interaction		
JOURNAL	Patent: US 6369212-A 1 09-APR-2002;		
FEATURES	Location/Qualifiers		
source	1..1781		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match 23.0%; Score 399.4; DB 2; Length 1781;			
Best Local Similarity 56.3%; Pred. No. 1e-112;			
Matches 826; Conservative 0; Mismatches 611; Indels 30; Gaps 3;			
Qy	99	AGAAACAGAAATCCACTCCAGGCGCATGGAAGTTTCTCATCATAGTAAATCTTCTCTCAAT	158
Db	86	AAAGCTAAACTTACTCTCTGTCACATGGAATTAACCTTTTATTTGAAGCCTACATCACT	145
Qy	159	TATTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGAC	218
Db	146	TGGCAGTGGCAGTCCACITCTCTCATCATGGCTAAAGAAATCTAGCAAACTTTATGGGC	205
Qy	219	CTGTTATGAGTCTTCAAAATTTGCCAAGTTTCAGCTGTTGTGTCATTTCTTCAGCTGAAGCAG	278

Db	206	CGCTCATGCACTTACGACTCGGGGAAATTCCTACCGTCATCATTTTCGTCCTCCGCGAATGG	265
Qy	279	CCAAAGAGGTTATGAAAACTCAGGCTGATGCTTTCGCCCAACGCGCTATCGTCTTTGGACG	338
Db	266	CGAAGGAAGTACTAAAACTCACGACCTCGCTTTTCGCAACGAGGCGGAAACTTGTGGTGG	325
Qy	339	CACAGATGTGTTTATAATCGGAAAGATGTCCTTGTGTTGCTTTCATATGGAGATCACTCGA	398
Db	326	CTGACATCGTCCATATATGATAGTAGTCGATATAGCAATTTTCTCCATATGTTGTAATACTCGA	385
Qy	399	GGCAGATGAAGAAATTTGGATCTTGAAATTTCTGAGTGCCAAAAAAGTTCAATCCTCCA	458
Db	386	GGCAGATTCGTAATAATTTTGCACTACGAACTCTCTTAGTGCCCAAGATGTCATAATCTTTA	445
Qy	459	GGTTAAATCCGAGAGGAAAGAAATGGAGGATGCCATCACAATTCCTCCGTTGCGAAAGCCGAT	518
Db	446	GCTCAATTCGCCAGATGAGCTGTCGATGATGGTCTCATCTATACGAACCATGCCAAATT	505
Qy	519	CTCCGGTCAATATTACAAAGATCAATTTATGCAATTAATTTTCGATCATGATAAGAACAT	578
Db	506	TTCCCGTCAACCTTACAGACAAAAATATTTTGGTTTACAAAGTTTCGGTAACTTGTAGATCAG	565
Qy	579	CCGTTGGTAA---TTGTAAGCAAAAGAAAGATGCTGAGTGTTCGCCGATGCAGTCAATG	635
Db	566	CTCTGGGAAAAAATATGTCGACCAAGACAACTGATTAATTTTCATGAGGAAATATAAT	625
Qy	636	AGGACGACGACGAGTTTGGCACCCGACGCTTTTCCGACGTGGAAATTTACTTCACTATA	695
Db	626	CATTGACAGGTGGATTTAGTATTGCTGATTTTCCCTACATGGAATGCTTACATGATG	685
Qy	696	TCATTGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACATATACTTTGAAG	755
Db	686	TTGTTGGTTCAAAAACCTAGACTGCTGTAAGGCTCATCGTAAAAATCGATGAGATTTTGGAA	745
Qy	756	AGATTCCTTAATGAACACAAAGCCAAATAGCCCT-----TTTG	791
Db	746	ATGTAGTGAATGAGCACAAACAGAAATCGAGCGGATGGCCAAAGGGTAATGCGGAATTTG	805
Qy	792	AAGCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTTCAGTGC	851
Db	806	CGGTTGAAGATTTGATCGATGTTTGTCTAAGGTTTCGAGAAAGTGGAGAAAGTTTCAAAAT	865
Qy	852	CAGTGACAAACGAAGAGCATCAAGCATCCGTTTTCGCAATGTTTACTCCCGGAGCGAAA	911
Db	866	CCATCACGGATGACAAATATCAAAATCAATATTAGTGGACATGTTCTCCGCTGGATCTGAAA	925
Qy	912	CAACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACATA	971
Db	926	CGTCATCGACCACTATAATTTGGGCATTTAGCTGAATGATGAAGAAACCAAGTGTCTTAG	985
Qy	972	GAAAGCAACAAGAAAGTTAGACAAGTATTGTTGGTGAATGGGAAAGTTGATGAATCA	1031
Db	986	CAAGGCCACAAGCTGAAGTGAGACAAGCTCTTGAAGGAAAGAAAGGTTTTCACAAATTTG	1045
Qy	1032	GATTTCAATGATTGAAATCTTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACATCCTC	1091
Db	1046	ATCTTGATGAGTTGAAGTACTTGAAGTTAGTAATCAAGAAACTCTAAGGATGCAACCTC	1105
Qy	1092	CGGT---TGTCTTGATTCGAGGGAGTGTAGAGAAACCAACGAATTTGATGATATGAAA	1148
Db	1106	CAATTCCTCTATTAGTCCCTAGAGAAATGATGAAGGATACAAAGATTGACGGGTACAATA	1165
Qy	1149	TTATCCGAAACACTCGAAATTTGTTGAATGCTTGGCGCATAGGAAGAGATCCTAATACTT	1208
Db	1166	TACCTTTCAAACTCGAGTCATAGTTAATGATGGGCAATTTGGACGAGATCCTGAAAGTT	1225
Qy	1209	GGTCCGAACCTCGAAAGTTTAAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTAAG	1268
Db	1226	GGGATGACCTCGAAAGCTTTTCCACAGAGAGATTCGAGAAATAGTTTCTGTGACTTCTTTG	1285
Qy	1269	GGACGACATTTGAACTGGTACCATTTGGTGCAGGAAAAAGAAATATGTCCTCGCATTTCTT	1328

Db 1286 GAAGCCATCATCAATTTATTCCTTTGGTGGGGAAGAGGATTTGCTCGGAATGCTTT 1345
QY 1329 CAGCTATTACCAATTTGGAGTATGTCATTATTAATCTATTATATCAATTTTAATTTGGGAC 1388
Db 1346 TTGGTTTAGCCAATGTTGGCAACCACTAGCTCAATTTACTTTATCACTTCGATCGGAAAC 1405
QY 1389 TGCCCGATGGAATTAACACCTCAACACCTTGATATGACTGAAGCTATTGGCGGTGCTCTCA 1448
Db 1406 TCCCTAATGACAAAGTCACGAAATTTGGACATGACGAGTACACCTGGAATTTCTGCAA 1465
QY 1449 GGAATAAATAGATCTTAAGTTGATTCCTTATTCATATCAATGAAGTTAGGCTGCAATA 1508
Db 1466 CRAAGAAAGGATGATCTGTTGTTGATTCGCCACCTTATGATCCTTGAATGATTGAGACA 1525
QY 1509 TTCTCTTGATTACATAGGAGGTTGAAA 1535
Db 1526 GTTGTAGAAATAAAGAGGGAGAA 1552

RESULT 9
AF122821 1781 bp mRNA linear PLN 08-FEB-2000
LOCUS Capsicum annum cytochrome P450 (PepCYP) mRNA, complete cds.
DEFINITION AF122821
ACCESSION AF122821
VERSION AF122821.1 GI:6739505
KEYWORDS
SOURCE Capsicum annum
ORGANISM Capsicum annum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Capsicum.

REFERENCE 1 (bases 1 to 1781)
Oh,B.J., Ko,M.K., Kim,Y.S., Kim,K.S., Kostenyuk,I. and Kee,H.K.
A cytochrome P450 gene is differentially expressed in compatible
and incompatible interactions between pepper (Capsicum annum) and
the anthracnose fungus, Colletotrichum gloeosporioides
Mol. Plant Microbe Interact. 12 (12), 1044-1052 (1999)
10624013

REFERENCE 2 (bases 1 to 1781)
Oh,B.J., Ko,M.K., Kostenyuk,I., Kim,Y.S., Kim,K.S. and Kee,H.K.
Direct Submission
Submitted (22-JAN-1999) Kumho Life and Environmental Science
Laboratory, Kumho Petrochemical Co., LTD., 1 Oryong-dong, Puk-gu.,
Kwangju 500-712, Korea

FEATURES
Location/Qualifiers
1..1781
/organism="Capsicum annum"
/mol_type="mRNA"
/db_xref="taxon:4072"
gene 1..1781
/gene="PepCyp"
CDS 4..1512
/gene="PepCyp"
/codon_start=1
/product="cytochrome P450"
/protein_id="AAF27282.1"
/db_xref="GI:6739506"
/translation="MEIQTNLVAFLFLSSILLKKWTKLNLPQGWKLPFGS
LHLAVAGLPFHGLKLNKXGPLMHLRGEIPVTIISPRMAKEVLKTHDLAFATR
PKLVADI VHYDSTDIAPGYEYWRQIRKICILELLSAKMYKFFSSIRQDLSMMVS
SIRTPNFVNITDKI FWFSTSVTCRSALGKIIIPMRIISITGFSIADF
PTMKLHDVGSKTRLLKAHRKIDILEHVNHEKONRADGKNGFEGGDLIDL
LVRSSGEVQISITDNIKSLVDMSAGSETSTIILWALAEMKKPSVLAKADEV
RQVLKEGQFQIDLDLKYLVIKETILRHPPILPLVPRECMKDKIDGYNIPKT
RVIVNAWALGRDPESWDPSPERSFNSVDFLGSHHQIFPFAGARRICPGLFGL
ANVGQPLAQLLYHFRKLPNGOSHENLDMTESPGISATRKDLDLVIATPYDP"

ORIGIN

Query Match 23.0%; Score 399.4; DB 4; Length 1781;
Best Local Similarity 56.3%; Pred. No. 1e-112;
Matches 826; Conservative 0; Mismatches 611; Indels 30; Gaps 3;
QY 99 AGAAACAGAAATCCACCTCCAGGGCCATGGAAGTTTCTCATCATAGGTAATCTTCTCTCAT 158

Db 86 AAAAGCTAAACTTACCTCTCTGTCATGGAATTAACCTTTATTTGAGAGCTACATCACT 145
QY 159 TATTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCTTTGGCTCAAAATTTATGGAC 218
Db 146 TGGCAGTGGCAGTCCACTTCTCATCATGCGCTAAAAAATCTAGCAAAATTTATGGGC 205
QY 219 CTGTTATAGTCTTCAAAATTTGCCAAGTTTTCAGCTGTTGTCATTTCTTACGTGGAAGCAG 278
Db 206 CGCTCATGCACTTACGACTCGGGGAAATTCCTACCGTCACTTTCGTCCCCGGAATGG 265
QY 279 CCAAGAGGTTATGAAAACTCAGGCTGATGCGCTTCGCCCAACGCCCTATCGTCTTGGACG 338
Db 266 CGAAGGAAGTACTAAAACTCAGACCTCGCTTTCGCAAGAGGCCGAAACTTGTGGTGG 325
QY 339 CACAGATTGTTTTATTAATCGGAAAGATGCTTTGTTGCTTCATATGAGAGATCACTGGA 398
Db 326 CTGACATCGTCCATTATGATAGTACGGAATATAGCAATTTCTCCATATGGTGAATACTGGA 385
QY 399 GGCAGATGAAGAAAATTTGGATACTTGAATTTCTGAGTCGCCAAAAAAGTTCAATCTCTCA 458
Db 386 GGCAGATTCTGTAATTTTGCATCTCGAACTCCTTAGTGCCAAAGATGGTCAAAATTTCTTTA 445
QY 459 GGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTCCTCCCTTCGAAAAGCCGGAT 518
Db 446 GCTCAATTCGCCAGGATGAGCTGTCGATGATGCTCATCTATACGAACTATGCAATTT 505
QY 519 CTCGGTCAATTAATCAAAAGATCAATTTATGGCATTAATAATTCGATCATGATGAAGACAT 578
Db 506 TTCCCGTCAACCTTACAGACAAAAATATTTTGGTTTACAAGTTCCGTAACCTTGTAGATCAG 565
QY 579 CGGTTGGTAA---TTGTAGCAAAAGAAAGATTCGCTGATGTTGCCGATGCTCAATG 635
Db 566 CTCGGGAAAAATATGCTGTGACCAGACAAAACCTGATAATTTTCATGAGGAAATTAATAT 625
QY 636 AGGCAGCCGACGAGTTTGGCAGCCGAGACGCTTTTCCGACGTGGAAATTAATTTCACTATA 695
Db 626 CATTGACAGGTGGAATTTAGTATTGCTGATTTTTCCTACATGGAATAATGCTACATGATG 685
QY 696 TCATTTGGAGCTGAGTCAAAACCCAGCGGTTTGCATCAGAGATTTGACGATATATCTTTGAAG 755
Db 686 TTGGTGGTTTCAAAAACTAGACTGCTGAAGGCTCATCGTAAAAATCGATGAGATTTTGGAA 745
QY 756 AGATTCTTAATCAACACAAAGCCAATAAGCCT-----TTTG 791
Db 746 ATGTAGTGAATGACACAAACAGAAATCGAGCGGATGGCCAAAGGGTAATGGGAAATTTG 805
QY 792 AAGCGGATAACTTAATGGATGTTCTATTGAAATCTTCAAAAAAATGGAAACGTTTCCAGTGC 851
Db 806 GCGGTGAAGATTTGATCGATGTTTTCCTAAGGGTTTCGAGAAAGTGGAGAAGTTCAAAATTT 865
QY 852 CAGTGACAAAACGAAAGCATCAAGCATCGGTTTTCGAAATGTTTACTGCGGGAGCGGAAA 911
Db 866 CCATCAGCGGATGACAAATATCAAAATCAATATTAAGTGACATGTTTCCCGCTGGATCTGAAA 925
QY 912 CAACCTTCGAAAGCTACAGAACTGGGTAATGGCAGAGCTGATGAAATAATCAACCTGAACTAA 971
Db 926 CGTCATCGACAATTAATTTGGGCAATTAGCTGAAATGATGAAGAAACCAAGTGTCTTAG 985
QY 972 GAAAAAGCACAAGAAGTTAGCAAGTATTTGGTGAATTTGGGAAATGGGAAAAGTTGATGAATCAA 1031
Db 986 CAAAGGCACAAGCTGAAGTGACAAAGTCTTCTAAGGAAAAGAAAGGTTTTTCAACAAATTTG 1045
QY 1032 GATTTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTTACATCCTC 1091
Db 1046 ATCTTGTAGTGTGAAGTACTTGAAGTTAGTAATCAAGAAACTCTTAAGGATGACACCTC 1105
QY 1092 CGGT---TGTCTTTGATTCGAGGGAGTGTAGAGAAACACACGAAATGATGATATGAAA 1148
Db 1106 CAATTCCTCTATTAGTCCCTAGAGAAATGTATGAAGGATCAAGATTTGACGGGTCAATA 1165
QY 1149 TTCTCCGAAACACTCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCCTTAATACTTT 1208

Db	1166	TACCTTTCAAAACTCGAGTCATAGTTAATGCAATGGGCAATTTGGACGAGATCCTCGAAAGTT	1225	Db	369	TACAATAGCTGTGATCTAGCCCTTTTGGCCCTATATGGGCACTATTGGAGACAAATGCGTAA	428	
Qy	1209	GGTCGGAACCTGGAAGTTTAAACCCAGAAAGGTTTAAAGATGTGTCAATTTGATATATAAG	1268	Qy	413	ATTTGGATPACTTGAATTTCTGAGTGCCAAAAAAGTTCAATCCTCCAGGTTAAATCCGAGAG	472	
Db	1226	GGGATGACCTGAAAGCTTTTCCCAGAGAGATTCGAGAAATAGTTCTGTGACTTTCTTG	1285	Db	429	ATATGTGTCTTGGAAAGTGCTCAGTGCCAAAGAAATGTTCCGACATTTAGCTCTATTATAGGCGG	488	
Qy	1269	GGACGACATTTGAACTGGTACCAATTTGGTGCAGGAAAAAGAAATATGTTCTCTGGCATTA	1328	Qy	473	GAAGAAATGGAGGATGCCATCACATTCCTCGTTTGGAAAGCCGATCTCCGGTCAATATT	532	
Db	1286	GAAAGCCATCATCAATTTATTCATTTGCTTGGTGGGGAAGAGGATTTGTCTCGAAATGCTTT	1345	Db	489	AATGAAGTTCTTTCGTCTCATTTAATTTTATCCGGTCATCTTCTGTGTGAACCTATTAAATG	548	
Qy	1329	CAGCTATTACCAATTTGGAGTATGTCATTTATAAATCTATTATATCATTTTAAATTTGGAA	1388	Qy	533	ACAAAGATCATTTTATGGCATTATAAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T	589	
Db	1346	TTGGTTAGCCAATGTTGGACAACCATTAGCTCAATTTACTTTTATCATTTCGATTCGAAAC	1405	Db	549	ACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG	608	
Qy	1389	TGGCCGATGGAATTAACACCTCAACACATTGATATGACTGAAGCTATTGGCGGTGCTCTCA	1448	Qy	590	TGTAAGCAAAAGAAAGATTGCTGAGTGTGGCGATGCGATCAATGAGGCGAGCGAGT	649	
Db	1406	TCCTAATGGCAAAAGTCAGAAATTTGGACATGACGGAGTCACCTGGAAATTTCTGCAA	1465	Db	609	TTCAAAGAGCAAGACAAATTTTATACAATAATTAAGAAGTGATATCTTTAGCAGGAGGG	668	
Qy	1449	GGAAAAAATAGATCTTAAGTTGATCTCTATTCATATCAATGTTAGCTTAGGCTCAATA	1508	Qy	650	TTTGGCACCGCAGAGCGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGGAGCTGAG	709	
Db	1466	CAAGAAGGATGATCTGTTTGTGATTCGACCCCTTATGATCCTTTGAATGTATTGAGACA	1525	Db	669	TTTATGATGGCTGACATATCCCTTCACTGAAGTTTCTTTCATGTGCTCAGTGGAAATGAAG	728	
Qy	1509	TTTCTTTGATTACATAGGAGGTTGAAA	1535	Qy	710	TCAAAACCCAGCGGTTTGCATCAGGAGATTGACGATATATCTTTGAAGAGATTCTTAAATGA	769	
Db	1526	GTGTAGAAATAAAAGAGGGGAGAAA	1552	Db	729	GGTAAGATTATGAATGCACACCATAGGTAGATGCCATTGTTGAGAATGTCAATCAATGAG	788	
RESULT 10				Qy	770	CAC-----AAAGCCAAATAAGCCTTTTGAAGCGGATCACTTAATG	808	
LOCUS	CS223966	1673 bp	DNA	linear	PAT 15-DEC-2005			848
DEFINITION	Sequence 222 from Patent WO2005111217.			Db	789	CACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGGAGCGTTTAGGAGGTGAAGATTTAAT		
ACCESSION	CS223966			Qy	809	GATGTTCTATTGAATCTTCAAAAAAATGGAACCGTTCCAGTGCCAGTCGACAAACGAAAGC	868	
VERSION	CS223966.1	GI:83685962		Db	849	GATGTTCTCTAAGACCTTATGAATGATGGAGGCCCTTCAATTTCTCATCACCACACGACAA	908	
KEYWORDS				Qy	869	ATCAAGCATCCGTTTGTGCAAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA	928	
SOURCE	Nicotiana tabacum (common tobacco)			Db	909	ATCAAGCTATAATTTTGTGACATGTTTGTCTGCCGGACGGAGACTTCATCGTCAACAAAT	968	
ORGANISM	Nicotiana tabacum			Qy	929	GAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCACAGAAAGAA	988	
REFERENCE	1			Db	969	GTGTGGGCTATGGTAGAAATGGTGAATAATCCAGCCGTTATTGCGGAAAGCTCAAGCAGAA	1028	
AUTHORS	Xu, D.			Qy	989	GTTAGACAAGTATTTGGTGAATGGGAAATGGGAAAGTTGATGAATCAAGATTTCATGATTGAAA	1048	
TITLE	Nicotiana nucleic acid molecules and uses thereof			Db	1029	GTAAGAGAGCATTTAGAGGAAAAAGAAACTTTCGATGAATAATGATGTGGAGGAGCTAAAC	1088	
JOURNAL	Patent: WO 2005111217-A 222 24-NOV-2005;			Qy	1049	TTCTTCAAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCTTGATT	1105	
FEATURES	U.S. Smokeless Tobacco Company (US)			Db	1089	TACCTAAAGTTAGTAATAAAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC	1148	
source	Location/Qualifiers			Qy	1106	CGAGGGAGGTGAGAGAAACCAACGAAATGATGATGAATTAATTCATCCGAACTCGA	1165	
ORIGIN	1..1673			Db	1149	CCAAAGAAATGTAGGAAAGAGACAAATATAAAACGGCTACACTATTTCCTGTAAGAACCAAA	1208	
	/organism="Nicotiana tabacum"			Qy	1166	ATTGTTGTGAATGCTTGGCGGATAGGAAGAGATCCTAATACTTGTGTCGGAACCTCGAAAG	1225	
	/mol_type="unassigned DNA"			Db	1209	GTCATGGTTAATGTTTGGGCTTTTGGGAAGAGATCCAAATAATTTGGAATGACCGAGAACT	1268	
	/db_xref="taxon:4097"			Qy	1226	TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAGGAGCGACATTTGAACTG	1285	
				Db	1269	TTTATGCCAGAGAGATTTGAGCAGTGTCTAAGGATTTTGTGGTAAATAATTTTGAATAT	1328	
				Qy	1286	GTAACATTTGGTGCAGGAAAAAGAAATATGCTCTCGCATTACTTTCAGCTATTACCAATTTG	1345	
				Db	1329	CTTCATTTGGTGGCGAAGGAGGATTTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT	1388	
				Qy	1346	GAGTATGTCATTATAAATCTATTATATCAATTTTAAATTTGGAACTGGCCGATGGAATTACA	1405	
				Db	1389	TATTTGCCATTGGCTCAATTACTATATCACTTCGATTGGAACTCCCTGCTGGAATCGAA	1448	
				Qy	1406	CCTCAAAACACATTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAATAGATCTT	1465	
				Db	1449	CAAAGCGACTTGGACTTGAAGTTGGTGGAGTAACTGCGCGCTAGAAAAAGTGACCTT	1508	

QY 1466 AAGTGTGATTCCTATTCATATCAA 1489
 Db 1509 TACTTGGTTCGGACTCCTTATCAA 1532

RESULT 11
 CS223950 1610 bp DNA linear PAT 15-DEC-2005
 LOCUS
 DEFINITION Sequence 206 from Patent WO2005111217.
 ACCESSION CS223950
 VERSION CS223950.1 GI:83685854
 KEYWORDS
 SOURCE Nicotiana tabacum (common tobacco)
 ORGANISM Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Nicotiana.
 1
 Xu, D.
 Nicotiana nucleic acid molecules and uses thereof
 TITLE Patent: WO 2005111217-A 206 24-NOV-2005;
 JOURNAL U.S. Smokeless Tobacco Company (US)
 FEATURES
 Location/Qualifiers
 1..1610
 /organism="Nicotiana tabacum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4097"

ORIGIN

Query Match 22.7%; Score 393.2; DB 2; Length 1610;
 Best Local Similarity 56.7%; Pred. No. 8.9e-111;
 Matches 825; Conservative 0; Mismatches 598; Indels 33; Gaps 4;

QY 101 AAACAGAAATCCACCTCCAGGCGCATGGAAGTTTCCTATCATAGGTAAATCTTCCTCATTTA 160
 Db 106 AAACTAAATTTGCTCTCTGTCATGGAATTAATACCTTTTATTTGGAAGTTTACACCAATTTG 165

QY 161 TTACTACATCTTGATCTAGGCGCATGAACGTTTATAGAGCTTGGCTCAAAATTTATGACCT 220
 Db 166 GCTGTGGAGGTCACCTTCTCACCATGGCCTAAATAATTTAGCCAAAGCTATGGTCT 225

QY 221 GTTATGAGCTTCAAATTTGGCCAAAGTTTCAGCTGTTGTCAATTTCTTTCAGCTGAAGCAGCC 280
 Db 226 CTTATGCAATTTCAACTTTGGCAAAATTTCTACACTGTCATATCATCATCACTCAAAATGSCA 285

QY 281 AAAGAGTTATGAAGAACTCAGGCTGATGCTTGGCCCAAGCCCTATCGTCTTTGACGCA 340
 Db 286 AAAGAAGTACTAAAGAACTCAGACCTCGCTTTTGGCCACTAGACCAAGCTTGTCTGGCC 345

QY 341 CAGATTGTGTTTTATAATCGGAAGATGCTTTGTTTGTCTTATATGAGATCACTGGAG 400
 Db 346 GACATCATCTACGACAGCAGGACATAGCACCTTTCGCCATACGGTGAATACTGGAGA 405

QY 401 CAGATGAAGAAATTTGGATCTTGAATTTCTGAGTGGCCAAAGAAAGTTTCAATCTCCAGG 460
 Db 406 CAAATTCGTAATTTTGCAATTTGGAACCTTTGAGTGGCCAAAGATGTCAGTTTATTAGC 465

QY 461 TTAATCCGAGAGAGAAATGAGGATGCCATCACATTCCTCGTTTCGAAAGCCGATCT 520
 Db 466 TCGATTCCGCAAGATGAGCTCTCGAAGATGGTTTCATCTATACGAACGACGCCCAATCTT 525

QY 521 CCGGTCAATATTACAAAGATCATTTATGGCATTATAATTTTCGATCATGATAGAACATCC 580
 Db 526 CCAATCATCTTACGACAGAGATTTTGGTTTACGAGTTCGGTAAATTTGTAGATCAGCT 585

QY 581 GTTGTAA---TTGTAAGCAAAAGAAAGATTTCTGAGTGTTCGCCGATGACGTCAATGAG 637
 Db 586 TTAGGGAAGATATGTGGTGACCAAGACAAATTTGATCATTTTATGAGGGAATAATATCA 645

QY 638 GCAGCGACGAGTTTGGCACCCGACGCTTTTCCGAGTGGAAATTTACTTCACTATATC 697
 Db 646 TTGGCAGGTGGATTTAGTATTGCTGATTTTTCCTTACATGGAATGATTCATGATATT 705

QY 698 ATTGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTTGAAGAG 757
 Db 706 GATGGTTCAAAATCTAAACTGGTGAAGGCACATCGTAAGATTGATGAATTTTGGAAAT 765

QY 758 ATTCTTAATGAACACAAA-----GCCAATAAGCCCTTTTGAA 793
 Db 766 GTGGTAATGAGCACACAAACAGAAATCGAGCAGATGGTAAAGGGTAAATGGTGAATTTGGT 825

QY 794 GCGGATACTTAATGGATGTTCTTATTTGAATCTTTCAAAAAAATGGAACGTTTCAGTGGCCA 853
 Db 826 GGAGAAGATCTGATTTGTTTGAAGATTAGAGAAAGTGGAGAAGTTCAATTTCCA 885

QY 854 GTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATCTTTTACTCCGCGGAGCGGAACA 913
 Db 886 ATCAGAGATGACAAATATCAATCAATTAATTCGACATGTTCTCTCCGGATCGGAACA 945

QY 914 ACTTCGAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAAATCAAACTGAACCTAAGA 973
 Db 946 TCATCGACAACTATAATTTGGGCAATTAGCTGAAATGATGAAGAACCAAGTGTTTTAGCA 1005

QY 974 AAAGCACAAGAAAGATTAGACAAGTATT---TGGTGAATGGGAAAAGTTGTATGATCA 1030
 Db 1006 AAGGCACAAGCTGAAAGTGAGCCAAGCTTTGAAGGGGGAAGAAAATTAGTTTTCAGAGATT 1065

QY 1031 AGATTTTCATGATTTCGAAATTTCTTCAAGTTAGTGGTTTAAAGAACTCTTAAGATTACATCCT 1090
 Db 1066 GATATTGATAGCTAAAGTATTGAGTTAGTGTATCAAGAAACTTTTAAAGATGCAACCT 1125

QY 1091 CCGGT---TGTCTTGATTCGAGGGAGTGTAGAGAAACACACGAATTTGATGATATGA 1147
 Db 1126 CCAATTCCTCTGTAGTCCCTAGAGAAATGATGGAAGATACAAAGATTGATGGTTACAAT 1185

QY 1148 ATTCATCGAACACTCGAATTTGTGTAATGCTTGGGCGATAGGAAGAGATCCTAATACT 1207
 Db 1186 ATACCTTTTCAAAACAAGAGTCAATTTGTAATGATGGCAATTTGGACGAGATCCTCAAAGT 1245

QY 1208 TGGTCGGAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTCGAATTCGATTTATA 1267
 Db 1246 TGGGATGATCTCGAAAGCTTTACGCCAGAGATTTGAGAAATATCTTATGATTTCTT 1305

QY 1268 GGGACGACATTTGAACTGTACCACTTTGGTGCAGGAAAAGAAATATGCTCGCATTACT 1327
 Db 1306 GGAATATCATCATCAATTTATTCATTTGGTGCAGGAAGAGATTGTCCTGGAATGCTA 1365

QY 1328 TCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGGAA 1387
 Db 1366 TTTGTTTGTAGCTAATTTGGACAACCTTTTAGCTCAGTTACTTTATCACTTCGATTGGAAA 1425

QY 1388 CTGGCCGATGGAATTTACACCTCAAACTTTGATATGACTTGAAGCTATTGGCGGTGCTCTC 1447
 Db 1426 CTCCTAATGGAACAACTACCAAAATTTTCGACATGACTGATGACCTGGAAATTTCTGCT 1485

QY 1448 AGGAAAAAATAGATCTTTAAGTTGATTCCTATTTCATATCAAGTTAGCTTAGGCTCAAAAT 1507
 Db 1486 ACAAGAAAGGATGATCTTATTTGATTGCACTCTGCTCATTTCTTGTATTAGTATTGCT 1545

QY 1508 ATTTCTTGATTACATA 1523
 Db 1546 GCTTTTCTATTGGAGA 1561

RESULT 12
 CS224014
 LOCUS
 DEFINITION Sequence 270 from Patent WO2005111217.
 ACCESSION CS224014
 VERSION CS224014.1 GI:83685886
 KEYWORDS
 SOURCE Nicotiana tabacum (common tobacco)
 ORGANISM Nicotiana tabacum

CS224014 Sequence 270 from Patent WO2005111217.
 CS224014.1 GI:83685886
 Nicotiana tabacum (common tobacco)
 Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

REFERENCE		asterids; lamiids; Solanales; Solanaceae; Nicotiana.
AUTHORS	1	
TITLE		Xu,D.
JOURNAL		Nicotiana nucleic acid molecules and uses thereof
FEATURES		Patent: WO 200511217-A 270 24-NOV-2005;
source		U.S. Smokeless Tobacco Company (US)
ORIGIN		Location/Qualifiers
Query Match		1..1610
Best Local Similarity		/organism="Nicotiana tabacum"
Matches		/mol_type="unassigned DNA"
		/db_xref="taxon:4097"
		22.7%; Score 393.2; DB 2; Length 1610;
		56.7%; Pred. No. 8.9e-111;
		0; Mismatches 598; Indels 33; Gaps 4;
Qy	101	AAACAGATCCACCTCCAGGCGCATGGAAGTTTCCTATCATAGGTAATCTTCCTCATTTA 160
Db	106	AAACTAAATTTGCGCTCCTGGTCCATGGAAATTTACCTTTTATTTGGAAGTTTACACCATTTG 165
Qy	161	TTACTCACTTCTGATCTAGGCGCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGAAGCT 220
Db	166	GCTGTGGCAGGTCCTCCTCACCATGCGCTAAATAATTTAGCCAAAGCTATGCTCT 225
Qy	221	GTATGAGTCTTCAAAATTTGGCCAAAGTTTCAGCTGTGTGTCATTTCTCAGCTGAAGCAGCC 280
Db	226	CTTATGTCATTTCAACTTGGCAAAATTTCTACACTCGTCATATCATCATCACTCAATCAATGGCA 285
Qy	281	AAAGAGTTATGAAACTCAGCTGATGCTTCGCGCAAGCCCTATCGTCTTTGAGCGCA 340
Db	286	AAAGAAGTACTAAAACTCACACCTCGCTTTTGGCCACTAGACCAAAAGCTTGTCTGTGCGCC 345
Qy	341	CAGATTGTGTTTTATAATCGGAAAGATGCTTGTGTTTCTTATATGAGATCACTCGGAGG 400
Db	346	GACATCATCTACAGCAGCAGCAGCATAGCACCTTCGCCATACGCTGAATCTGGAGA 405
Qy	401	CAGATGAAGAAAATTTGGATCTTGAAATTTCTGAGTGGCCAAAAGTTCAATCTCCAGG 460
Db	406	CAAAATCGTAAAATTTGTCATATTTGAACTCTTTGAGTGGCCAAAGATGGTCAAGTTTTTATAGC 465
Qy	461	TTAATCCGAGAGAAAGATGAGATGCCATCACATTCCTCGTTTCGAAAGCCGATCT 520
Db	466	TCGATTGCGCAAGATGAGCTCTCGAAGATGGTTTCATCTATACGAAACGACGCCCAATCTT 525
Qy	521	CCGGTCAATATTACAAAGATCAATTTATGGCATTATAATTTTCGATCATGATAGAAACATCC 580
Db	526	CCAGTCAATCTTACCGACAAGATTTTGTGTTTACGAGTTGGTAAATTTGTAGATCAGCT 585
Qy	581	GTGGTAA---TTGTAAGCAAAAAGAAAGATTGCTGAGTTGGCCGATGCAGTCAATGAG 637
Db	586	TTAGGAAGATATGTGGTGACCAAGACAAATTTGATCAATTTTATGAGGGAATTAATATCA 645
Qy	638	CGACGCGAGTTTTGGCACCCGACGCTTTTCCGAGTGGAAATTTACTTCACTATATC 697
Db	646	TTGGCAGGTGGATTTAGTATGCTGATTTTTCCTTACATGAAATAGATTCTGATAT 705
Qy	698	ATTGAGCTGAGTCAAAACCCAGGCGTTTTCATCAGGAGATTGACGATATACTTTGAAGAG 757
Db	706	GATGGTTCAAAATCTAAACTGGTGAAGGCACATCGTAAGATTGATGAAATTTTGGAAAT 765
Qy	759	ATCTTTAATGAACACAAA-----GCCAATAGCCTTTTGA 793
Db	766	GTGGTAAATGAGCAAAAACAGAATCGAGCAGATGGTAAAGGGTAAATGGTAAATTTGGT 825
Qy	794	GCGGATACTTAATGATGCTTCTATTGATCTTCAAAAAATTTGAAACCTTCAGTGGCA 853
Db	826	GGAGAAGATCTGATTGATGTTTTGTTAAGAGTTAGAGAAAGTGGAGAGTTCAAAATCCA 885
Qy	854	GTGCAAAACGAAAGCATCAAGCATCCGTTTTTGCAAAATGTTTACTGCCGGGACGGAACA 913
Db	886	ATCACAGATGACAATATCAATCAATATTAAATCGACATGTTCTCTCGCGGATCGGAAACA 945
Qy	914	ACTTCGAAAGCTACAGAAATGGGTAATGCGCAGAGCTGATGAAAAATCCAACTGAACCTAAGA 973
Db	946	TCATCGCAACTATAATTTTGGGCATTAGCTGAAATGATGAAGAAACCAAGTGTGTTTAGCA 1005
Qy	974	AAAGCACAAGAAGATTAGACAAGTATT---TGGTGAATGGGAAAAAGTTTGATGAATCA 1030
Db	1006	AAGGCACAGCTGAAGTGAGCCAGCTTTGAAGGGGAAGAAATTTAGTTTTCAAGAGATT 1065
Qy	1031	AGATTTTCATGATTTGAAATTTCTCAAGTTAGTGTAAAGAAACTCTTAAGATTACATCT 1090
Db	1066	GATATTGATAAGCTAAAGTATTTCGAAGTTAGTATCAAAAGAAACTTTTAAAGAAATGCACCT 1125
Qy	1091	CCGGT---TGTCTTGATTCGAGGGAGTGTAGAGAAACAACACGAAATTTGATGGATATGAA 1147
Db	1126	CCAAATTCCTCTGTTAGTCCCTAGAGAAATGATATGGAAGATACAAAGATTGATGGTTTACAAT 1185
Qy	1148	ATTCAATCGAACACTCGAAATTTGTTGAATTTGTTGGCGGATAGGAAGAGATCCCTAATACT 1207
Db	1186	ATACCTTTCAAAACAAGAGTCAATTTGTTAATGCTATGGGCAATTTGGACGAGATCCTCAAAGT 1245
Qy	1208	TGGTCGGAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAATAA 1267
Db	1246	TGGATGATCCTGAAAGCTTTTACGCCAGAGAGATTTGAGAAATAATTTCTATTGATTTTCTT 1305
Qy	1268	GGGACGACATTTGAACTGGTACCATTTTGGTGCAGGAAAAAGAAATATGTCCTGGCATTACT 1327
Db	1306	GGAAATCATCATCAATTTATTCCATTTTGGTGCAGGAAGAGATTGTCCTGGAAATGCTA 1365
Qy	1328	TCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCAATTTTAATTTGGAA 1387
Db	1366	TTTGGTTAGCTAATTTTGGACAACCTTTAGCTCAGTTTACTTATCACTTCGATTGGAAA 1425
Qy	1388	CTGCGCGATGAAATTAACACTCAACCTCAACACTTGATGATGACTGAAGCTATTGGCGGCTCTC 1447
Db	1426	CTCCCTAATGACAAACTCACAAAATTTTCACATGACTGAGTCACTGGAATTTTCTGCT 1485
Qy	1448	AGGAAAAAATAGATCTTAAGTTGATTCCTTATCCATATCAAGTTAGCTTAGGCTCAAAAT 1507
Db	1486	ACAAGAAAGGATGATCTTATTGTTGATGCCACTCTCTGCTCATCTTGTATTGAATTGCT 1545
Qy	1508	ATTTCTTGATTACATA 1523
Db	1546	GCTTTCTATTGGAGA 1561
RESULT 13		
LOCUS		CS223944 1566 bp DNA linear PAT 15-DEC-2005
DEFINITION		Sequence 200 from Patent WO200511217.
ACCESSION		CS223944
VERSION		CS223944.1 GI:83685851
KEYWORDS		
SOURCE		Nicotiana tabacum (common tobacco)
ORGANISM		Nicotiana tabacum
AUTHORS		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
JOURNAL		asterids; lamiids; Solanales; Solanaceae; Nicotiana.
FEATURES		1
source		Xu,D.
		Nicotiana nucleic acid molecules and uses thereof
		Patent: WO 200511217-A 200 24-NOV-2005;
		U.S. Smokeless Tobacco Company (US)
		Location/Qualifiers
		1..1566
		/organism="Nicotiana tabacum"
		/mol_type="unassigned DNA"
		/db_xref="taxon:4097"
ORIGIN		
Query Match		22.4%; Score 388.8; DB 2; Length 1566;
Best Local Similarity		56.5%; Pred. No. 2.1e-109;
Matches		801; Conservative 0; Mismatches 587; Indels 30; Gaps 3;

QY 101 AAACAGAAATCCACTCCAGGCGCATGGAAGTTTCTTATCATAGTAATCTTCTCATTTA 160
DB 98 AAACATAATTTGGCTCTGCTCATGGAATTAACCTTTTATTGGAAGTTTACACCAATTG 157
QY 161 TTACTCATTCTGATCTAGGCCATGAAGTTTATAGAGCCTTGCTCAAAATTTATGACCT 220
DB 158 GCTGTGGCAGGCTCACTTCTCACCATTGCGCTTAAATAATTTAGCCAAAGCTATGCTCT 217
QY 221 GTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAAGCTGTTGTCATTTCTTCAAGCTAAGCAGC 280
DB 218 CTTATGCAATTTACAACTTTGGCAAAATCTTACACTCATATCATCACTCAATATGGCA 277
QY 281 AAAGAGGTTATGAAATCTCAGGCTGATGCTTGGCCCAACGCCCTATCGTCTTTGGACGCA 340
DB 278 AAAGAAGTACTAAAACTCAGACCTCGCTTTTGGCCACTAGACCAAAGCTTGTGCTGGCC 337
QY 341 CAGATTGTCTTTATTAATCGGAAGATGCTTGTGTTGCTTCTATATGAGATCACTGGAGG 400
DB 338 GACATCAITTCATACACAGCAGGACATAGCATTTTCTCCGTACGGTGAATCTGGAGA 397
QY 401 CAGATTGAAGAAATTTGGGATCTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCCAGG 460
DB 398 CAAATTCGTAAATTTGCATATTTGGAACCTTGGAGTCCCAAGATGGTCAAAATTTTATAG 457
QY 461 TTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTTCTCGTTTCGAAAGCCGATCT 520
DB 458 TCGATTTCGCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGACGACACCCAATCTT 517
QY 521 CCGGTCAATATTACAAAGATCAATTTATGGCATTTAATTTTCGATCATGATAGAAACATCC 580
DB 518 ACAGTCAATCTTACTGACAAAAATTTTGTGTTTACGAGTTCGGTAACTTGTAGATCAGCT 577
QY 581 GTTGTGTA--TTGTAAAGCAAAAGAAAGATGCTGAGTGTCCGATCAGTCAATGAG 637
DB 578 TTAGGAAGATATGTTGACCAAGACAATTTGATCAITTTTATAGGGAATAATATCA 637
QY 638 CGACGACGAGTTTGGCACCGCAGACGCTTTTCGACGTGGAAATTTACTTCACTATATC 697
DB 638 TTGGCAGGTGGATTAGTATTCTGATTTTTCCTTACATGGAATGATTCATGATTT 697
QY 698 ATTGAGCTGATCAAAACCCAGGCGTTGATCAGGAGATTGACGATATCTTGAAGAG 757
DB 698 GATGTTTGAATCTAATCTGTTGTAAGCAACATCTGTAAGATTGATGAAATTTTGGGAAT 757
QY 758 ATTCTTAATGAACAAAA-----GCCAATAAGCTTTTGA 793
DB 758 GTTGTGTGACACAAAAAGAACAGACAGACGATGGCAAGAGGTTAATGGTGAATTTGT 817
QY 794 GCGGATACTTAATGGATTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTGCCA 853
DB 818 GGTGAAGATTTGATTGATGATTTGTTAAGAGTTAGAGAAAGTGGAGAAAGTTTCAAAATTCCT 877
QY 854 GTGACAAAGCAAGCATCAAGCATCCGTTTGGCAATTTGTTTACCTGCGGGAGCGAAACA 913
DB 878 ATCAAAATGCAAAATCAATCAATTAATTAATCGACATGTTCTCTGCGGGATCTGAAACA 937
QY 914 ACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAATCAAGA 973
DB 938 TCATCGACGACTTAATTTGGGCATTTAGCTGAATGATGAAGAAACCAAGTTTGTAGCA 997
QY 974 AAAGCAAGAAAGATTAGACAAGTATTTGGTGAATTTGGGAAAAAGTTGATGAATCAAGA 1033
DB 998 AAGGCACAAGCTGAAGTAAGGCAAGCTTTGAAGGAGAAAAAGGTTTCAACAGATTGAT 1057
QY 1034 TTTCAATGATTTGAAATCTTCAAGTTAGTTGTTTAAAGAAAATCTTAAGATTTACCTCGG 1093
DB 1058 CTTGATGAGCTAAAAATATCTCAAGTTAGTAAATCAAGAAACCTTTAAGAAATGCAACCTCCA 1117
QY 1094 GT---TGCTTCTGATTTCCGAGGAGGTAGAGAAACAACACGAATTTGATGGATATCAAT 1150
DB 1118 ATTCCTCTATTAGTTCTTAGAATTTGATGGAGGATACAAGATTGATGGTTTACAAATATA 1177
QY 1151 CATCCGAACACTCGAATTTGTGTAATGCTTGGGCGATAGGAAGAGATCTTAATACTTTGG 1210

DB 1178 CTTTTCAAAACAAGAGTCATAGTTAATGATGGCAATCGACGAGATCCAGAAAGTTGG 1237
QY 1211 TCGGAACCTTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAAGGG 1270
DB 1238 GATGACCCCGAAAGCTTTATGCGCAGAGAGATTGAGATAGTTCTATTGACTTTCTTGA 1297
QY 1271 ACGACATTTGAACCTGTTACCATTTGGTGAGGAAAGAAATATGCTCGCATTAATTCTCA 1330
DB 1298 AATCATCATCAGTTTATACCATTTGGTGAGGAAAGGATTTGTCCGGAATGCTATTT 1357
QY 1331 GCTATTACCAATTTGGAGTATGTCATTTATAAATCTATTATATCAITTTTAAATTCGGAACTG 1390
DB 1358 GGTTTAGCTAATTTGGACAACTTTAGTCTAGTTACTTTATCACTTCGATTTGGAACCTC 1417
QY 1391 GCGGATGGAATTTACACTCAACACTTTGATATGATGATGAGGCTATTTGGCGGCTCTCAGG 1450
DB 1418 CCTAATGGCAAAAGTCATGAGAAATTTGACATGACTGACTGACCTGGAATTTCTGCTACA 1477
QY 1451 AAAAAAATAGATCTTAAAGTTGATTCCTATTCCATATCA 1488
DB 1478 AGAAGGATGATCTTGTGTTTGTGCTCTCTTATGA 1515

RESULT 14
CS223902
LOCUS CS223902 1576 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 158 from Patent WO2005111217.
ACCESSION CS223902
VERSION CS223902.1 GI:83685432
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1
AUTHORS Xu D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 200511217-A 158 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
LOCATION/Qualifiers
FEATURES
source 1..1576
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="caxon:4097"

ORIGIN

Query Match 22.4%; Score 388.8; DB 2; Length 1576;
Best Local Similarity 56.5%; Pred. No. 2.1e-109;
Matches 801; Conservative 0; Mismatches 587; Indels 30; Gaps 3;

QY 101 AAACAGAAATCCACTCCAGGCGCATGGAAGTTTCTTATCATAGTAATCTTCTCATTTA 160
DB 110 AAACATAATTTGGCTCTGCTCATGGAATTAACCTTTTATTGGAAGTTTACACCAATTG 169
QY 161 TTACTCATTCTGATCTAGGCCATGAAGTTTATAGAGCCTTGCTCAAAATTTATGACCT 220
DB 170 GCTGTGGCAGGCTCACTTCTCACCATTGCGCTTAAATAATTTAGCCAAAGCTATGCTCT 229
QY 221 GTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAAGCTGTTGTCATTTCTTCAAGCTAAGCAGC 280
DB 230 CTTATGCAATTTACAACTTTGGCAAAATCTTACACTCATATCATCACTCAATATGGCA 289
QY 281 AAAGAGGTTATGAAATCTCAGGCTGATGCTTGGCCCAACGCCCTATCGTCTTTGGACGCA 340
DB 290 AAAGAAGTACTAAAACTCAGACCTCGCTTTTGGCCACTAGACCAAAGCTTGTGCGGGCC 349
QY 341 CAGATTGTCTTTATTAATCGGAAGATGCTTGTGTTGCTTCTATATGAGATCACTGGAGG 400
DB 350 GACATCTTACTACGACGACGACATAGCATTTTCTCCGTACGGTGAATTAATGACTGAGA 409
QY 401 CAGATGAAGAAAAATTTGGATACTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCTCCAGG 460

Db 410 CAAATTCGTAATAATTTGCATATTTGGAACCTTTGAGTGCACAAGATGGTCAAAATTTTTTAGC 469
Qy 461 TTAATCCGAGGAGGAAGAAATGAGGATGCCATCACATTCCTCCGTTTCGAAAGCCGATCT 520
Db 470 TCGATTCGCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAACACACCCCAATCTT 529
Qy 521 CCGGTCAATATTTACAAAGATCATTTATGGCAATTAATAATTTGATCATCATGATAAGAACATCC 580
Db 530 ACAGTCAATCTTACTGACAAAATTTTTTGGTTTACGAGTTCGGTAACCTTGATGATCAGCT 589
Qy 581 GTTGATAA---TTGTAAGCAAAAAGAAAGATGCTGAGTGTTCGCGATGCAGTCAATGAG 637
Db 590 TTAGGGAAGATATGTTGGTACCAAGCAAAATTTGATCATTTTTATGAGGGAATAATATCA 649
Qy 638 GCAGCGACGAGTTTTGGCACCGCACAGCGCTTTTCGAGCTGGAATTTACTTCACATATATC 697
Db 650 TTGGCAGGTGGATTTAGTATGCTGATTTTTTCCCTACATGGAATAATGATTCATGATAT 709
Qy 698 ATTGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTTGAAGAG 757
Db 710 GATGCTCGAATCTAAATCTGATGGAAGCACATCTGAAGATTTGATGAATTTTGGGAAT 769
Qy 758 ATCTTATGAACACAAA-----GCGAATAAGCCTTTTGA 793
Db 770 GTTGTGATGACACAAAAGAACAGACGATGCAAGAGGGTAATGGTGAATTTGGT 829
Qy 794 GGGGATAACTTAATGGATGTTCTATTTGAATCTTTCAAAAAATGGAACGTTCCAGTGC 853
Db 830 GGTGAAGATTTGATTTGATTTGTTAGAGTTAGAGAAAGTGGAGAGTTCAATTCCT 889
Qy 854 GTGCAAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGCAACA 913
Db 890 ATCACAATGACAATATCAATCAATATTAATCGACATGTTCTCTGCGAGGATCTGAAACA 949
Qy 914 ACTTCGAAGCTACAGATGGTGAATGCGAGAGCTGATGAAATAATCCAACTGAATAGA 973
Db 950 TCATCGACGACTAATTTTGGGCAATTTAGCTGAATTTAGTGAATAACCAAGTGTTTTGA 1009
Qy 974 AAGACACAAGAAAGTAGTAGCAAGTATTTGTGAAATGGGAAAAAGTTGATGAATCAAGA 1033
Db 1010 AAGGCACAAGCTGAATGAGGCAAGCTTTGAAGGAGAAAAAGGTTTTCACAGATTGAT 1069
Qy 1034 TTTTCATGATTTGAAATTTCTCAAGTTAGTGGTTTAAAGAACTCTTAAGATTTACATCCTCG 1093
Db 1070 CTTGATGAGCTAAAAATATCTCAAGTTAGTAAATCAAAAGAAACCTTAAGAATGCAACCTCCA 1129
Qy 1094 GT---TGTCTGATTTCCGAGGAGTGTAGAGAAACACACGAATTTGATGCGATGAAAT 1150
Db 1130 ATTCCTCTATTAGTTCTTAGAGAAATGATGAGGAGTACAAAGTTGATGGTTTACAAATA 1189
Qy 1151 CATCGAACACCTCGAATTTGTTGAATGCTTCGGCGATAGGAAGAGATCTTAATACTTGG 1210
Db 1190 CTTTTCNAAACAAGAGTCATAGTTAATGATGGGCAATCGGACGAGATCCAGAAAGTTGG 1249
Qy 1211 TCGGAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTTGCAATTTGATTAATAAGGG 1270
Db 1250 GATGACCCCGAAAGCTTTATGCGCAGAGATTTGAGAAATAGTCTTATTGACTTTCTTGG 1309
Qy 1271 ACACATTTGAACTGGTACCATTTTGGTCAGGAAAAAGAAATATGCTTGGCAATTAATCTCA 1330
Db 1310 AATCATCATCAGTTTATACCATTTGGTCAGGAAGAGGATTTGTCGGGGAATGCTATTT 1369
Qy 1331 GCTATTACCAATTTGGAGTATGTCATTAATAAATCTATTATATCAATTTTAAATTTGGGA 1390
Db 1370 GGTTTAGCTAATGTTGGACAACCTTTAGCTCAGTTTACTTTATCACTTCGATTTGGAAC 1429
Qy 1391 GCCGATGGAAATTAACCTCAACACATTTGATATGATGAGAGCTATTGGGGGCTCTCAGG 1450
Db 1430 CCTAATGGCAAAAGTCATGAGAATTTTCGACATGACTGAGTCACTCGGAATTTCTGCTACA 1489
Qy 1451 AAAAAAATAGATCTTAAGTTGATTTCTATTTCCATATCA 1488

Db 1490 AGAAAGGATGATCTTGTGTTTGTGATTTGCCACTCCTTATGA 1527

RESULT 15
CS223904
LOCUS 1581 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 160 from Patent WO2005111217.
ACCESSION CS223904
VERSION CS223904.1 GI:83685433
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

1
Xu,D.
Nicotiana nucleic acid molecules and uses thereof
Patent: WO 2005111217-A 160 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
Location/Qualifiers
1. .1581
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

FEATURES
source
ORIGIN

Query Match 22.4%; Score 398.8; DB 2; Length 1581;
Best Local Similarity 56.5%; Pred. No. 2.1e-109;
Matches 801; Conservative 0; Mismatches 587; Indels 30; Gaps 3;

Qy 101 AAACAGAAATCCACCTCCAGGSCCATGAGATTTTCTATCATAGGTAATCTTCCTCATTTA 160
Db 115 AAATCAAAATTTGCCCTCCCTGCTCATGGAATTTACCTTTTATTTGGAAGTTTACACCATTTG 174
Qy 161 TTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCTTTGGCTTCAAAATTTATGAGACT 220
Db 175 GCTGGGAGGTGCATCTTCCTCACCGGCTTAAATAATTTAGCCAAAGCTATGCTCT 234
Qy 221 GTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAGCTGTGTGTCATTTCTTACGTGGAAGC 280
Db 235 CTTATGCAATTTACAACTTGGACAAATCTTCACTCATCATATCATCACCTCAAAATGGCA 294
Qy 281 AAAGAGTTATGAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGCA 340
Db 295 AAAGAAGTACTAAATAACTCACGACCTGCTTTTGCCACTAGACCAAAAGCTGTGCTGGCC 354
Qy 341 CAGATTGTTGTTTATAATCGGAAGATGCTTTGTTTGTCTTATATGAGATCACTGGAGG 400
Db 355 GACATCATCTACTAGCAGCAGCAGCATAGCATTTTCTCCGTACGGTGAATACTGGAGA 414
Qy 401 CAGATGAAGAAAAATTTGGATCTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGG 460
Db 415 CAAATTCGTAATAATTTGCAATTTGGAACCTTTGAGTGCACAAAGATGGTCAAAATTTTTTAGC 474
Qy 461 TTAATCCGAGAGGAAGAAATGGAGATGCCATCACATTCCTCCGTTCCGAAGCCGATCT 520
Db 475 TCGATTCCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAACACACCCCAATCTT 534
Qy 521 CCGGTCAATATTTACAAAGATCATTTATGGCAATTAATAATTTGATCATGATAAGAACATCC 580
Db 535 ACAGTCAATCTTACTGACAAAATTTTTTGGTTTACAGTTCCGTAACCTTGATAGACGT 594
Qy 581 GTTGATAA---TTGTAAGCAAAAAGAAAGATGCTGAGTGTTCGCGATGCAGTCAATGAG 637
Db 595 TTAGGGAAGATATGTTGGTGAACCAAGACAAATTTGATCATTTTTATGAGGGAATAATATCA 654
Qy 638 GCAGCGACGATTTTGGCACCGCAGACGCTTTTCGAGCTGGAATTTACTTCACATATATC 697
Db 655 TTGGCAGGTGATTTAGTATGCTGATTTTTTCCCTACATGGAATAATGATTCATGATAT 714
Qy 698 ATTGGAGCTGAGTCAAAACCCAGGCGTTTGTGATCATGAGGATTTGACGATATATCTTGAAGAG 757

Db 715 GATGGTTCGAATCTAAACTGGTGAAGCAGACATCGTAGATTGATGAAATTTTGGGAAT 774
Qy |||||
Db 758 ATTCTTAATGAACACAAA-----GCCAATAGCCTTTTGAA 793
Db 775 GTTGTGATGAGCACAAGAAAGAACAGAGCAGATGGCAAGAGGTAAATGGTGAATTTGGT 834
Qy 794 GCGGATACTTAATGAGTCTCTATTGAATCTTCAAAAAATCGAAACGTTCCAGTGCCA 853
Db 835 GGTGAAGATTTGATTGATGTATTGTTAAGAGTTAGAGAAAGTGGAGAAGTTCAATTCCT 894
Qy 854 GTGACAAACGAAGCATAAGCATCCGTTTTCGAAATGTTTACCTGCGCGGAGCGAAACA 913
Db 895 ATCACAATGACAATATCAATCAATTAATCGACATGTTCTCTGCGGATCTGAAACA 954
Qy 914 ACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAATGA 973
Db 955 TCATCGACGACATTAATTTGGGCATTAGCTGAATGATGAAGAAACCAAGTGTTTTAGCA 1014
Qy 974 AAAGCACAAGAGAAGTTAGACAAGTATTGGTGAAATGGGAAAAAGTTGATGAATCAAGA 1033
Db 1015 AAGGCACAAGCTGAAGTAAGGCAAGCTTTGAAGGAGAAAAAGGTTTTCACACAGATTGAT 1074
Qy 1034 TTTCTAGATTTGAAATCTTCAAGTTAGTGGTTAAAGAAACCTCAAGATTACCTCCG 1093
Db 1075 CTTGATGAGCTAAATATCTCAAGTTAGTAAATCAAGAAACCTTAAGAAATGCACCTCCA 1134
Qy 1094 GT---TGTCTTGATTCGAGGAGGTGTAGAGAAACAAACAGAAATTGATGATATGAAATT 1150
Db 1135 ATCTCTATTAGTTCTTAGAAGATGTATGGAGGATACAAAGATTGATGGTTACAATATA 1194
Qy 1151 CATCCGAACACTCGAATTTGTTGAATGCTTGGCGGATAGGAGAGATCCTAATACTTGG 1210
Db 1195 CCTTTCAAAAACAAGAGTCATAGTAAATGATGGCAATCGGACGAGATCCAGAAAGTTGG 1254
Qy 1211 TCGGAACCTGGAAGTTTAAACCAGAAAGTTTAAAGATTGCGAATGATTAAGGG 1270
Db 1255 GATGACCCCGAAAGCTTTATGCCAGAGAGATTGAGAAATGTTCTATTGACTTCTTGGGA 1314
Qy 1271 ACGACATTTGAACCTGTTACCAATTTGGTGCAGGAAAGAAATATGCTGGCAATTACTTCA 1330
Db 1315 AATCATCATCAGTTTATACCAATTTGGTCAGGAAGAGATTGTCGGGAAATGCTATTT 1374
Qy 1331 GCTATTACCAATTTGGAGTATGTCATTATATAATCTATTATATCAATTTAAATGGGAACGT 1390
Db 1375 GGTTTAGCTAATTTGGACAACCTTTAGCTCAGTTACTTTATCACTTCGATTGGAACTC 1434
Qy 1391 GCCGATGGNAATTACACCTCAACACTCTGATATGATGAAGCTATTGGCGGTGCTCTCAGG 1450
Db 1435 CCTAATGGACAAGTCATGAGAATTTGACATGACTGAGTCACTGGAATTTCTGCTACA 1494
Qy 1451 AAAAAAATAGATCTTAAGTTGATTCCTATTCCATATCA 1488
Db 1495 AGAAAGGATGATCTTTGTTGATGGCACTCCTTATGA 1532

Search completed: May 31, 2006, 11:50:40
Job time : 10159 sec8

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 09:01:43 ; Search time 8717 Seconds
(without alignments)
11117.151 Million cell updates/sec

Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gcataaaggaaatggagc.....tcttttcaatccgaaaaa 1733

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est3:
3: gb_est4:
4: gb_est5:
5: gb_est6:
6: gb_hnc:
7: gb_est2:
8: gb_est7:
9: gb_est8:
10: gb_est9:
11: gb_gss1:
12: gb_gss2:
13: gb_gss3:
14: gb_gss4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371.2	21.4	654	10	DV154738
2	366.2	21.1	732	10	DV153215
3	347.2	20.0	696	10	DV150019
4	342	19.7	766	10	DV134278
5	336.4	19.4	710	10	DV131796
6	332.8	19.2	542	10	DV137266
7	304.6	17.6	781	10	DV137008
8	294.6	17.0	793	9	DN487910
9	293	16.9	778	10	DV145886
10	286.4	16.5	681	10	DV138221
11	275.2	15.9	965	10	DF506382
12	274.2	15.8	706	10	DV119846
13	273.2	15.8	902	8	CX053093
14	271.8	15.7	674	10	DV128282
15	270	15.6	819	9	CX636845
16	270	15.6	965	9	CX670312
17	266.8	15.4	933	9	CX636844
18	261.6	15.1	794	10	DV124816
19	261	15.1	718	10	DV122991

20	258.8	14.9	742	10	DV148280
21	255.8	14.8	849	9	CX669008
22	255.8	14.8	900	9	CX663419
23	254.8	14.7	903	9	CX668839
24	253.8	14.6	731	10	DV126517
25	252.2	14.6	614	10	DV147676
26	252.2	14.6	872	10	DM003125
27	247.8	14.3	675	10	DV143418
28	247.8	14.3	799	10	DV127937
29	247.6	14.3	709	4	CA921009
30	245.2	14.1	742	10	DV153269
31	245.2	14.1	1016	9	CX701446
32	245	14.1	730	4	CA922231
33	243.6	14.1	738	10	DV136837
34	242.2	14.0	780	10	DV149902
35	242.2	14.0	834	10	DT730556
36	241.8	14.0	790	10	DV697990
37	240.8	13.9	902	8	CO125331
38	239.6	13.8	830	10	DR925887
39	237	13.7	692	10	DV144641
40	237	13.7	720	1	AI730111
41	236.8	13.7	599	7	BE204783
42	236.8	13.7	689	2	BM063831
43	236.8	13.7	994	9	CX666031
44	235.4	13.6	711	4	CA921435
45	235	13.6	918	10	DR933028

ALIGNMENTS

RESULT 1
DV154738
LOCUS
DEFINITION
CV03095B2F04.f1 CV03-normalized library Euphorbia esula cDNA clone
DV154738 654 bp mRNA linear EST 03-OCT-2005
DV154738.1 GI:76863745
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Euphorbia esula (leafy spurge)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbieae; Euphorbia.
1 (bases 1 to 654)
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector) . . .TAAGCTTGATATCG(End Vector) (Start
EcoRI adaptor)AATTCATGTTGTGGG (End EcoRI adaptor) (Start
Insert) . . .AAAAAAAAAAAAAAAA(End Insert) (Start Tag)TCGCT (End
Tag) (Start NotI site/Vector)GGCGCCGCCGCGG . . . Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector trimming: Cross_match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG(T7)
BACKWARD: ATTAAACCTCACTAAAG(T3)
Insert Length: 654 Std Error: 0.00
Plate: CV03095B2 row: F column: 04
Seq primer: TAATACGACTCACTATAGG (T7)

High quality sequence stop: 654.
Location/Qualifiers
1..654
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03095B2F04.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcoRI(5' side of insert); Site 2: NotI(3' side of insert); Tissue: Seasonal crown buds from outdoor plants, whole plant tissue (leaf, stem, flowers, seeds, shoots, galls, root and crown buds) from outdoor plants, cold acclimated tissues (leaf, stem, and crown buds), crown and root buds from decapitated greenhouse p ants, and seedlings. The library was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. An identifying tag was added at the 3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGGT."

ORIGIN

Query Match 21.4%; Score 371.2; DB 10; Length 654;
Best Local Similarity 75.8%; Pred. No. 3.4e-84;
Matches 473; Conservative 0; Mismatches 148; Indels 3; Gaps 1;

QY 57 TTCTGCTGTTTAAATCTTAGTAGTATGATGAGGTGTGGAGAAACAGAAATCCACCTC 116
DB 31 TTCTAATATCTTCACTTATTAAAGTTAAACATAGAACATGACTCAAAATGTCCTC 90
QY 117 CAGGCCCATGGAAGTTCCTATCATAGGTAATCTTCCTCATTTATCTACTCTCTGATC 176
DB 91 CAGGCCCATGGAAGTTCCTATCATAGGTAATCTTCCTCATTTATCTACTCTCTGATC 150
QY 177 TAGGCCATGAAGCTTTAGAGCTTGGCTCAATTTATGGACCTGTATGACTTCAAA 236
DB 151 AAATCATCAAGCTTTAGCTCTTATCTCTAATTTATGGACCTATTATGAGTCTTCAAC 210
QY 237 TTGGCCAAAGTTTCAGCTGTGTGTCATTTCTTCAAGCTGAAGCAGCCAAAGAGGTTATGA 296
DB 211 TCGGCCAAGTTCCTAATTTATGTCATTTCTTCGGCTCAAGCAGTTAAAGAGGTTATGA 270
QY 297 CTCAGCTGATGCTTTCGCCCAACGCCCTATCGTCTTGGACCGCAGATGTGTTTTATA 356
DB 271 CTCATGTTGAATCTTCGCGAACGTTCTCCGGCGCTAGATCGCAGATTGTTTTATG 330
QY 357 ATCGGAAGATGCTGTGTTGCTTCATATGGAGATCACTGGAGGCGAGATGAAGAAATTT 416
DB 331 ATAGGAATGATATCTTTTGTGTCATATGGAAATCATTTGGAGGCAATGAGAAATTT 390
QY 417 GGATACCTTGAATTTCTGAGTGCCAAAGAAAGTTCAATCTCTCCAGGTTAATCCGAGAGGA 476
DB 391 GGATACCTGAATTTTATGTCATAGCAGATTCAATCTTTCAGGTTAATCCGGAGGA 450
QY 477 AAATGAGATGCGATCAATCTTCCTCGTTTCAAGCCCGGATCTCCGGTCAATATACAA 536
DB 451 AAGTTTCGGAGGCGATCAATCTTCCTCGTTTCAAGAAAGTGGATCTCCGGTCAATCTTAC 510
QY 537 AGATCAATTTATGGCATTAATTTTCGATCATGTAAGACATCCGTTGCT---AATTGTA 593
DB 511 AAGTCATATTTGAGTTTCAAAATTTCAATCATGATGAAGCAATCATTTGTAAGAAATGTC 570
QY 594 AGCAAAAGAAAGATTTGCTGAGTTGTCGGATGCGAGTCAATGAGGCGAGCGAGTTTG 653
DB 571 ACCAAAGAAAGACTTATGATGATTGTTGGCGGTGTGAATGAGGCAATTTTCGAGTTTC 630
QY 654 GCACCGCAGACGCTTTTCGAGCT 677
DB 631 GGATTCGGATGCTTTCCCGTCT 654

RESULT 2
DV153215

LOCUS DV153215 732 bp mRNA linear EST 03-OCT-2005
DEFINITION CV03091A2E05.f1 CV03-normalized library Euphorbia esula cDNA clone
CV03091A2E05.f1 5, mRNA sequence.
ACCESSION DV153215
VERSION DV153215
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiaceae; Euphorbia.
1 (Bases 1 to 732)
REFERENCE Anderson,J.V., Horvath,D.P., Thimmapuram,J., Liu,L., Hernandez,A.,
Kim,W.R. and Mikel,M.
Direct submission, Anderson,J.V. 2005
JOURNAL Unpublished (2005)
COMMENT Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector) . . . TAAGCTTGATCG(End Vector) (Start
EcoRI adaptor)/AATTCATTGTGTGGG (End EcoRI adaptor) (Start
Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TCGGT (End
Tag) (Start NotI site/Vector) GCGGCCGCCACCGCG . . . Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector trimming: Crose_match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATCAGCTCACTATAGG(T7)
BACKWARD: ATTAACCTCACTAAG(T3)
Insert Length: 732 Std Error: 0.00
Plate: CV03091A2 row: E column: 05
Seq primer: TAATCAGCTCACTATAGG (T7)
High quality sequence stop: 732.
Location/Qualifiers
1..732
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03091A2E05.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcoRI(5' side of
insert); Site 2: NotI(3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGGT."

FEATURES
source

ORIGIN

Query Match 21.1%; Score 366.2; DB 10; Length 732;
Best Local Similarity 70.3%; Pred. No. 6.8e-83;
Matches 507; Conservative 0; Mismatches 208; Indels 6; Gaps 1;

QY 587 AATTCATAGCAAAAGAAAGATTGCTGAGTGTTCGGATGCAATGAGCAGCGAGC 646
DB 12 AAATGCAAGAATCAAGAGCGGTGATGATGTGATTGATTCGAGTGAAGCTGAGCT 71
QY 647 AGTTTGGCAGCGCAGACGCTTTTCGACGCTGGAAATTTACTTCACTATATCAATTGAGCT 706
DB 72 GGGTTCAGTGTTCGCGACGATTTTCGCTCGCTAAATTTCTTCACTATCTGAGTGAGAA 131

QY	707	GAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATACCTTGAAGAGATTTCTTAAT	766
Db	132	AAGTCAAAGTTGCAGAAAGTTGCATAAAGACACAGATGAGATTCCTTGAAGAGATTATAAGT	191
QY	767	GAACACAAAGCCCAATAGCCCTTTTG-----AAGCGATCACTTAATGGATGTTCTATTG	820
Db	192	GAACATAAGACTTAATGCTAAGATTTGGAAAGTCAAGCTGATTAATCTTTTGGATGTTTGTG	251
QY	821	AATCTTCAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAAGCATCAAGACATCC	880
Db	252	TATCTTCAGAAAATGGAATCTTCAAGTTCCATTTGACTAATGATATATCAAGACGCT	311
QY	881	GTTTTCGAAATGTTACTGCGGAGCGAAACAACTTCGAAAGCTACAGAAATGGGTAATG	940
Db	312	ACTCTGGAATGTTTCGGAGCTGGGAGCGACATCTCTCAAAAACCTACAGACTGGCAATG	371
QY	941	GCAGAGCTGATGAAAAATCAACTGAAGTAAGAAAAGCACAAGAAAGTTAGACAAGTA	1000
Db	372	CGCAACTAATGAGGAAGCCATCAACATGAAAAAGGCCCAAGAGGTTAGCGGGTC	431
QY	1001	TTTGGTGAAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTA	1060
Db	432	TTTGGCGATACGGGAAGGTAGAGAAATCAAGAAATCCAGAACTGAAATTACTTGAATTA	491
QY	1061	GTGGTTAAAGAAACTCTAAGATTACATCTCTCCGGTTGCTTTGATTCGAGGGAGTGTAGA	1120
Db	492	GTGTTAAAGAAACATTTGAGATTACATCTCTGCGGTGATTTCTTCTAGAGATGCCGA	551
QY	1121	GAACAAACAGAAATGATGATGAAATTCATCCGAACTCGAATGTTGTGAATGCT	1180
Db	552	GAGAAAACTAAAAATTCACGGATATGATGTTTATCTTAAAAAATAATTTCTGTGAATCT	611
QY	1181	TGGCGCATAGGAAGAGATCTTAATCTTGTGTCGAACTCGAAAGTTTACCCAGAAAGG	1240
Db	612	TGGCGAATTTGGAAGATCCCAAGTTTGACCGACCCGGAAGTTTCAACCCAGAAAGA	671
QY	1241	TTTAAAGATTGTGCAATTTGATTAATAAGGACGACATTTGAACTGGTGATTCATTTGGTGA	1300
Db	672	TTTCAAGATAGCCCAATAGACTACAAAGGTACAAATTTGCAATTTGATTCGTTGGTGT	731
QY	1301	G 1301	
Db	732	G 732	
RESULT 3			
DV150019			
LOCUS			
DEFINITION	DV150019 696 bp mRNA linear EST 03-OCT-2005		
ACCESSION	CV03081BIH01.fl CV03-normalized library Euphorbia esula cDNA clone		
VERSION	CV03081BIH01.fl 5, mRNA sequence.		
KEYWORDS	DV150019.1 GI:76859026		
SOURCE	EST.		
ORGANISM	Euphorbia esula (leafy spurge)		
	Euphorbia esula		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
	rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbiales;		
	Euphorbiaceae; Euphorbia.		
REFERENCE	1 (bases 1 to 696)		
AUTHORS	Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,		
	Kim, W.R. and Mikel, M.		
TITLE	Direct submission, Anderson, J.V. 2005		
JOURNAL	Unpublished (2005)		
COMMENT	Contact: James V. Anderson		
	USDA/ARS		
	1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND		
	58105, USA		
	Tel: 701 239 1263		
	Fax: 701 239 1252		
	Email: andersjv@fargo.ars.usda.gov		
	Adaptors and tags in 5'-end sequenced clones:		

(Vector) . . . TAAGCTTGATATCG (End Vector) (Start	
EcoRI adaptor) AATTCCTAGTTGTTGGG (End EcoRI adaptor) (Start	
Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TGCCT (End	
Tag) (Start NotI site) (Vector) GCGCGCCGACCGCGG. Base	
Calling/Quality Scores: PHRED from Washington University Genome	
Center. Vector Trimming: Cross match from Washington University	
Genome Center PHRAP suite. Low quality bases (Phred score < 20)	
were trimmed from both ends of the sequence by an in-house script.	
This sequence is vector free and at least 200 bp in length.	
PCR Primers	
FORWARD: TAATACGACTCACTATAGG (T7)	
BACKWARD: ATTAAACCTCACTAAAG (T3)	
Insert Length: 696 Std Error: 0.00	
Plate: CV03081BI row: H column: 01	
Seq primer: TAATACGACTCACTATAGG (T7)	
High quality sequence stop: 696.	
FEATURES	
Location/Qualifiers	
1. .696	
/organism="Euphorbia esula"	
/mol_type="mRNA"	
/db_xref="taxon:3993"	
/clone="CV03081BIH01.fl"	
/lab_host="DH10B"	
/clone_lib="CV03-normalized library"	
/notes="Vector: pBS II SK(+); Site 1: EcoRI (5' side of	
insert); Site 2: NotI (3' side of insert); Tissue:	
Seasonal crown buds from outdoor plants, whole plant	
tissue (leaf, stem, flowers, seeds, shoots, galls, root	
and crown buds) from outdoor plants, cold acclimated	
tissues (leaf, stem, and crown buds), crown and root buds	
from decapitated greenhouse plants, and seedlings. The	
library was constructed and normalized as described by	
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996). Genome	
Research 6 (9): 791-806. An identifying tag was added at the	
3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGCT."	
ORIGIN	
Query Match 20.0%; Score 347.2; DB 10; Length 696;	
Best Local Similarity 69.0%; Pred. No. 5.3e-78;	
Matches 475; Conservative 0; Mismatches 213; Indels 0; Gaps 0;	
QY 831 AAATGGAACGTTCCAGTGCAGTCACAAACGAAAGCATCAAAGCATCCGTTTTGCAAA 890	
Db 1 AACATGGAGATCTTCAACTTCGTTGACTGATGAAGTCATCAAGGCCAATGTTATGGAAA 60	
QY 891 TGTTTACTGCCGGAGGAAACAACTTCGAAAGCTACAGATGGTAAATGGCAGACTGA 950	
Db 61 TATTTTTTGGCGGAGTCATACATCTTCCATACTGTGGAATGGGCTATGTCGGAGTTGA 120	
QY 951 TGAATAATCCAACCTAAAGAAAGCAAGAAAGTTAGACAAGTATTTGGTGAAA 1010	
Db 121 TGAAGGAGCCACATCAATGAGNAAGACAGAGAGGTGAGACGAGTCTTTGGCAGA 180	
QY 1011 TGGGAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTAGTGGTAAAG 1070	
Db 181 AAGGAAAAGTCGAGGAGTCAAGAAATTCAGAAATTAACCTCATGTCAATTTGTCAAAG 240	
QY 1071 AAATCTTAGATTACATCTCTCGGTTGCTTGTGATTCGAGGAGGTAGAGAAACACAC 1130	
Db 241 AAATCTTAGATTACATCTCTCGGTTGCTTGTGATTCGAGGAGGTAGAGAAACACAC 300	
QY 1131 GAATTCATCGATATGAAATTCATCCGAAACACTCGAATTTGTTGTAATGCTTGGCGCATAG 1190	
Db 301 AGTCAATGATATGATATTCACCCAGACTGTACTTTCATGTTATGATATGGGCAATCG 360	
QY 1191 GAAGAGATCCTAATACTTGTGTCGGAACCTCGGAAGTTTAAACCCAGAAAGGTTTAAAGATT 1250	
Db 361 GAAGAGACCCCGATGTTTGGATTGATGCGGAAAAGTTTATCCAGAAAGATTGGAAGATG 420	
QY 1251 GTGCAATGATATTAAGGAGCAGACATTTGAACTGGTACCATTTTGTGTCAGGAAAAGAA 1310	
Db 421 GTCAATAGATTACAAAGGTGCAAAATATGGAACATAATCCGTTTGGTTCAGGAAAAGAA 480	

SOURCE
ORGANISM
Euphorbia esula (leafy spurge)
Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiae; Euphorbia.
1 (bases 1 to 710)
REFERENCE
ANDERSON, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
TITLE
JOURNAL
COMMENT
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector) . . . TAAGCTTGATATCG(End Vector)(Start
EcoRI adaptor)AATTCATGTTGGG(End EcoRI adaptor)(Start Tag)TGGGT(End
Insert) . . . AAAAAAAAAAAAAA(End Insert)(Start Tag)TGGGT(End
Tag)(Start NotI site/Vector)GGCGCCGCCCGG . . . Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector Trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCATATAGG(T7)
BACKWARD: ATTAACCTCTAAAG(T3)
Insert Length: 710 Std Error: 0.00
Plate: CV03061A1 row: F column: 07
Seq primer: TAATACGACTCATATAGG (T7)
High quality sequence stop: 710.
Location/Qualifiers
1. .710
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03061A1F07.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcoRI (5' side of
insert); Site 2: NotI (3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAATGCGT."

Query Match 19.4%; Score 336.4; DB 10; Length 710;
Best Local Similarity 67.6%; Pred. No. 3.2e-75;
Matches 472; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 660 CAGACGCTTTCCGACGTGGAATTAATCTTCACTATATATGAGCTGAGTCAAAACCCA 719
|||||
Db 13 CAGATCTGTTCTCTTTGAGATTTCTTAAGTATATTAATGGTGAANAATCTAGATGG 72
|||||

QY 720 GCGTTTGCATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCA 779
|||||
Db 73 AGAGGGTTCAAAAAGAGTGGATAAGATGCTTGATGATATTTCTTAAAGAACATAAAGTC 132
|||||

QY 780 ATAAGCTTTTGAAGCGGATACTTAATGGATGTTCTATGATCTTCAAAAAATGGAA 839
|||||
Db 133 AACGCAAGTATGAGGATAATAATCTTTTACAGTTTTTATGGATGTACAAGAATCGGAG 192
|||||

QY 840 ACGTTCCAGTGCAGTGACAAAGCAATCAAGCATCCGTTTTCGCAATGTTTACTCG 899
|||||

Db 193 ATCTTCAACTTCGGTTGACTGATGAAGTCATCAAGCCCAATGTTATGGAATATTTTG 252
|||||

QY 900 CCGGAGCGAAACAATTCGAAGCTACAGAATGGTAATGGCAGAGCTGATCAAAAATC 959
|||||

Db 253 GCGGAGCCATACATCTTCCAAAACCGTGAATGGCTATGTCGAGTTGATGAGGAGCC 312
|||||

QY 960 CAACTGAATTAAGAAAGCACAGAGAGAGTTAGACAAGTATTGTTGTAATGGGAAAG 1019
|||||

Db 313 CACAGTCAATGAGAAAGCACAGAGAGGTGAGACGAGTGTTCGAGAAAGGAAAG 372
|||||

QY 1020 TTGATGAATCAAGATTTTCATGATTTTGAATTTCTCAAGTTAGTGGTTAAAGAACTCTAA 1079
|||||

Db 373 TCGAGGAGTCAAGAATTCAGAAATTCGAATACCTCATGTCGATGTTCAGAAATCTGA 432
|||||

QY 1080 GATTACATCCTCGGTTGCTTGAATTCGAGGAGGTGTAGAGAAACAAACAGATGATG 1139
|||||

Db 433 GATTACACCCCTCCATCTACATTCGTTACGAGAGAAATGCAGAGTAAAAAACAAAAGTCAATG 492
|||||

QY 1140 GATATGAATTCATCCGAACACTCGAATTTGTGTAATGCTTGGCGGATAGGAGAGATC 1199
|||||

Db 493 GATATGAATTTACCCCAAGACTGTACTTCAATGTTAATGTGCGCAATCGGAGAGACC 552
|||||

QY 1200 CTAATACTTGGTGGAACTCGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGCAATTG 1259
|||||

Db 553 CCGAGTTTGGATGATGCTGAAGTTTATCCAGAAAGATTTGAAGACAGTCAAAATAG 612
|||||

QY 1260 ATTATAAGGAGCGACATTTGAACCTGGTACCAATTTGGTGCAGAGAAAGAAATATGCTGT 1319
|||||

Db 613 ATTACAAAGGTGCAATATGGAACTAATACCGTTTGGTGCAGAGAAAGAAATATGTCGG 672
|||||

QY 1320 GCATTACTTCAGCTATTACCAATTTGGAGTATGTCATT 1357
|||||

Db 673 GAATCACTTTCGCCATATCTTATGTGAGACTTCTCCTT 710
|||||

RESULT 6
DVI37266 542 bp mRNA linear EST 03-OCT-2005
LOCUS
CV03116B1H09.f1 CV03-normalized library Euphorbia esula cDNA clone
DEFINITION
CV03116B1H09.f1 5, mRNA sequence.
DVI37266
DVI37266.1 GI:76846272
EST.
Euphorbia esula (leafy spurge)
ORGANISM
Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiae; Euphorbia.
1 (bases 1 to 542)
ANDERSON, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
TITLE
JOURNAL
COMMENT
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector) . . . TAAGCTTGATATCG(End Vector)(Start
EcoRI adaptor)AATTCATGTTGGG(End EcoRI adaptor)(Start Tag)TGGGT(End
Insert) . . . AAAAAAAAAAAAAA(End Insert)(Start Tag)TGGGT(End
Tag)(Start NotI site/Vector)GGCGCCGCCCGG . . . Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector Trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers


```
Db 148 ATGGCGGTGACTGCCATTGGTAAAGAATATGGGATCAAGAAGCTTTGTTAAAGTATAATT 207
QY 623 GATGCAGTCAATGAGCGAGCAGCTTTTGGCACCGCAGACGCTTTTCGCGCTGGAA 682
Db 208 GATGTTGGAGTTGAGGGAGTGGCGGTATAGCATTTGCTGACGTGTTCTTCCTTTAAA 267
QY 683 TTACTTCACTATATCAATGGAGCTGAGTCAAAACCCAGGCGTTTGTGCATCAGGAGATTGAC 742
Db 268 TTCTCTTCAATACATACCGGAGAAATCTCGAACTGATCCGGTTGCGATGAAGACGGAC 327
QY 743 GATATACCTTGAAGAGATCTTAATGAACACAAAGCAATAAG----- 784
Db 328 AAGTACTCTTGAAGAGATCTTAACCGAAAGTAAAGAGAGAGAGAGAGAGTGGAGATCGTCGT 387
QY 785 CTTTTTGAAGCGGATAAATTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAACGTT 844
Db 388 GGTGGCGAAATCGAAATCTTTTGGATGTTCTTTTGGATTTTCCAGAAATGGAATCTT 447
QY 845 CCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTGTGCAAAATGTTTACTGCCGGG 904
Db 448 CCAGTCCCGTTGACGAATGCTAGCATCAAGGAGCCATTCTGGAATGTTTGGAGCGGG 507
QY 905 AGCGAAACAACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCAACT 964
Db 508 AGTGACACATCTCTCCAAAGACTACAGAAATGGGCAATGGCGCAACTAATGAGGCATCCAATA 567
QY 965 GAACTAAGAAAGACCAAGCAAGGTTAGACAAGTATTGTTGCAATGGGAAAGTTGAT 1024
Db 568 GAAATGAGAAAGTGCAGAGAGGTGAGCGAGTCTTTTGGTGAACCGGAAACGTTGAG 627
QY 1025 GAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTA 1084
Db 628 GAATCAAGGCTTGATGAACCTGACTACTTTGAAAGTTGTTGAAAGAAACGCTGAGACTA 687
QY 1085 CATCTCCGGTGTCTTGATTCGAGGAGGTAGAGAAACAAACAGCAATTTGATGATAT 1144
Db 688 CATCTCCGGGT-CATTGATTCCTAGAGNATCGGTGAGAGAGAAATCGGTGGATAT 746
QY 1145 GAAATTCATCCGAACACTCGAATTTGTTGTAATG 1178
Db 747 GATATTCATCCCAAGACTAAAGTATCGTGAACG 780
```

```
RESULT 8
LOCUS DN487910/c 793 bp mRNA linear EST 10-MAR-2005
DEFINITION R014G04.3pr Populus root cDNA library Populus tremula x Populus
tremuloides cDNA clone R014G04 3', mRNA sequence.
ACCESSION DN487910
VERSION 1
KEYWORDS EST.
SOURCE Populus tremula x Populus tremuloides
ORGANISM Populus tremula x Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 793)
AUTHORS Sterky,F., Bhalerao,R.R., Unneberg,P., Segerman,B., Nilsson,P.,
Brumer,A.M., Charbonnel-Campaa,L., Lindvall,J.J., Tandre,K.,
Strauss,S.H., Sundberg,B., Gustafsson,P., Uhlen,M., Bhalerao,R.P.,
Nilsson,O., Sandberg,G., Karlsson,J., Lundberg,J. and Jansson,S.
A Populus EST resource for plant functional genomics
Proc. Natl. Acad. Sci. U.S.A. 101 (38), 13951-13956 (2004)
15353603
Other_ESTs: R014G04, R014G04.5pr
Contact: Bo Segerman
Umea Plant Science Center, Department of Plant Physiology
Umea University
901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: bo.segerman@plantphys.umu.se.
Location/Qualifiers
```

FEATURES

source

```
1. .793
/organism="Populus tremula x Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:47664"
/clone="R014G04"
/tissue_type="root"
/clone_lib="Populus root cDNA library"
```

ORIGIN

```
Query Match 17.0%; Score 294.6; DB 9; Length 793;
Best Local Similarity 65.6%; Pred. No. 28-64;
Matches 460; Conservative 0; Mismatches 239; Indels 2; Gaps 2;
QY 790 TGAAGCGGATACTTAATGGATGTTCTATTGTAATCTTCA-AAAAATGGAAGCTTCCAG 848
Db 775 TGAAGCGGAGGATCTTCTGGATGTTCTTTTGAATCTCCAGAAAAACACGGGACCTTTGGAT 716
QY 849 TGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCCGGGAGCG 908
Db 715 TCCTTTTGACACACAGACAGCATCAAGCAACAATCTCTGGAATCTATTCGCGCTGGAGTG 656
QY 909 AAACAACCTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCCAACCTGAAC 968
Db 655 ATACATCATCAACATTAAATGGAATGGACAATGGCCGAAATGTTTAGAAATCCGATCGTAA 596
QY 969 TAAGAAAGGACCAAGCAAGGTTAGACAAGTATTGTTGCAATGGGAAAGTTGATGAAT 1028
Db 595 TCGAAAGGACACAGAGGAAGTGAGGCAAGTCTTTAGTAATATCTAGAAATGTTGATGAAA 536
QY 1029 CAAGATTTTCATGATTTGAAATCTTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTTACATC 1088
Db 535 CATGCTTTCACAATTTTGAATTTCTTGAAGCTGATTTATCAAGAAACTCTGAGACTGCATC 476
QY 1089 CTCGCGTTGCTTGTGATTCGAGGGAGGTAGAGAAACAAACAGCAATTTGATGATATGAAA 1148
Db 475 CTCGAGCTCCTTTGATTTCCAGAGAAATGCAACAAGACGTCGAGATTAATGGATATGTCA 416
QY 1149 TTCTATCCGAAACACTCGAATTTGTTGAAATGCTTTGGCGGATAGGAAGAGATCTTAATCTT 1208
Db 415 TACAGGCTCAAAAGTAAAGTATGATCAATCGTGGGCTATTGGAAGAGATTTCTGATCATTT 356
QY 1209 GGTCCGAAACCTGGAAAGTTTAAACCAGAAAGGTTTTAAAGATTTGTGCAATTTGATTAAG 1268
Db 355 GGACTGGAAGCTGAGAAATTTCTATCCAGAGAGATTTCTAGACAGTTTCGATTTGATTAAG 296
QY 1269 GGACGACATTTGAACTGGTACCATTTGTTGGTCAAGGAAAAAGAAATATGTCCTGCAATTTCT 1328
Db 295 GTACTTAATTTTGAATTTTCCCATTTGGTCTGGGAGAGGATGTTCTCTGCAATTTTAT 236
QY 1329 CAGCTATTACCAATTTGGAGTATGTCATTATAAAATCTATTATATCATTTTAAATTCGGAAC 1388
Db 235 TTGGTATAGCTACTGTCGAGCTTCCACTTTGGCAGCTGCTATACCATTTTCGATTTGAAAC 176
QY 1389 TGGCCGATGGAATTAACACTCAACACTTGTATGATGACTGAAGCTATTGGCGGTGCTCTCA 1448
Db 175 TTCCAAACCGAGACCTTTTGGAGATCTCGACATGAATGAAAGTTTTTGGGGCACGGTTA 116
QY 1449 GAAAAAATAAGATCTTA-AGTTGATTTCTATTTCATATCA 1488
Db 115 GAAGGAAACATCAACTTACAGCTGATTTCCCATTTCCCTTTTA 75
```

RESULT 9

```
DV145886
LOCUS DV145886 778 bp mRNA linear EST 03-OCT-2005
DEFINITION CV03068A1A02.f1 CV03-normalized library Euphorbia esula cDNA clone
CV03068A1A02.f1 5, mRNA sequence.
ACCESSION DV145886
VERSION DV145886.1 GI:76854893
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae; Euphorbieae; Euphorbia.

1 (bases 1 to 778)

Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A., Kim, W.R. and Mikel, M.

Direct submission, Anderson, J.V. 2005

Unpublished (2005)

Contact: James V. Anderson

USDA/ARS

1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND

58105, USA

Tel: 701 239 1263

Fax: 701 239 1252

Email: andersjv@fargo.ars.usda.gov

Adaptors and tags in 5'-end sequenced clones:

(Vector). . . TAGCTTGATATCG (End Vector) (Start
EcoRI adaptor) AATCCATTTGTTGGG (End EcoRI adaptor) (Start
Tag) (Start NotI site/Vector) GCGCGCCACCGCGG. . . Base
Insert). . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TCGGT (End
Tag) (Start NotI site/Vector) GCGCGCCACCGCGG. . . Base
Calling/Quality Scores: PHRED from Washington University
Center. Vector Trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAATACGACTCACTATAGG (T7)

BACKWARD: ATTAACCCCTCACTAAAG (T3)

Insert Length: 778 Std Error: 0.00

Plate: CV03068A1 row: A column: 02

Seq primer: TAATACGACTCACTATAGG (T7)

High quality sequence stop: 778.

Location/Qualifiers

1..778

/organism="Euphorbia esula"

/mol_type="mRNA"

/db_xref="taxon:3993"

/clone="CV03068A1A02.f1"

/lab_host="DH10B"

/clone_lib="CV03-normalized library"

/note="Vector: pBS II SK(+); Site 1: EcoRI (5' side of
insert); Site 2: NotI (3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonald, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAAATCGGT."

FEATURES

source

ORIGIN

Query Match 16.9%; Score 293; DB 10; Length 778;
Best Local Similarity 66.1%; Pred. No. 5.1e-64;
Matches 457; Conservative 0; Mismatches 225; Indels 9; Gaps 2;
QY 107 AATCACCTCCAGGCGCATGGAAGTTTCTATCATAGTAAATCTTCCTCATTTATTACTC 166
DB 66 ATCTCTCCCGCGGACCATGGAATTTCCCTCATCGGAACATCTCTCA-----AGTC 119
QY 167 ACTCTGTATCTAGGCGCATGAACGTTTTAGAGCGTTGGCTTCAAAATTTATGGACCTGTATG 226
DB 120 GCGGAGCTCTTCCACATCACCAGATTACAGAGATTTAGCTTAAATAATATGGGCTGTATG 179
QY 227 AGCTCTCAAAATTTGGCCAGTTTTCAGCTGTGTTCATTCTTCAGCTGAAGCAGCAAGAG 286
DB 180 AGTATTCAACTCGGCCAAGTTTCCATCCGTCGTCGTTTCATCAGCTGAAACTGCTAAAGAA 239
QY 287 GTTATGAAACTCAGGCTCATGCTTCCGCCCACACGCCCTTATCGTCTTGGACGCACAGATT 346
DB 240 GTCTCAAAACTCAGACGCTCCAGTTCGCGCGCGCGCGGCGGCTGCTCTGCGCGCACAAATT 299

QY 347 GTGTTTTATATCGAAAGATGCTCTTGTTCCTTCATATGAGATCATCTGAGGCGAGATG 406
DB 300 GTGTTTTATATCGATGATATTTGTGTTTGGTGAATATGAGATCTCTGGAGGCAAAATG 359
QY 407 AAGAAATTTGGATCTACTTGAATTTCTGAGTGCCTCCAAAAAAGTTCAATCTCCAGGTTAATC 466
DB 360 AGGAAATTTTGCATCTTTAGAAATTAATCTAGTGCCTCCAAAAAAGTTCAATCTTTAGACCGGTT 419
QY 467 CGAGAGGAAGAAATGAGGATGCCATCACATTCCTCCGTTCCGAAAGCGGATCTCCGGTC 526
DB 420 CGGAAGCAGAAAGTTGCGAATTTTGTGATTTTCTCCGGTCCGAAAGCGGAAATCACTT 479
QY 527 AATATTACAAAGATCATTTATGGCATTAATTTTCGATCATGATGAAGAACATCCGTTGGT 586
DB 480 AATCTTACGAAACTATATTTTCGTTTAAACGAATTCGATCATGGCGGTGACATGCGCAATGGT 539
QY 587 ---AATTGTAAGCAAAAGAAAGATTGCTGAGTGTTCGCGATGCAATCAATGAGGCGAGCG 643
DB 540 AAGAAATATGGGATCAGAAAGCTTTGTAAGTATTAATTTGTTGGAGTTGAGGAGGT 599
QY 644 ACAGATTTTGGCACCGGACGACGCTTTTCCGACGTTGGAAATTTACTTCACTATATCAATTGGA 703
DB 600 GCGGTTTATAGCAATGCTGACGTTGTTCTTCTTAAATTCCTTCAATACATCAGGGA 659
QY 704 GCTGAGTCAAAACCCAGCGGTTTGCATCAGGAGATGACGATATATATTTGAAGAGATTCTT 763
DB 660 GAATACCTCGAAACTGATCCGGTTGCATAAGAAAGACAGACAAAGTACTTTGAAGAGATCTTA 719
QY 764 AATGAACACAAAGCCCAATTAAGCCTTTTGAAG 794
DB 720 AGCAAGTAAGAGAGAGAGAGAGAGTGGAG 750

RESULT 10

DV138221
LOCUS DV138221 681 bp mRNA linear EST 03-OCT-2005
DEFINITION CV03119A2G06.f1 CV03-normalized library Euphorbia esula cDNA clone
CV03119A2G06.f1 5, mRNA sequence.
ACCESSION DV138221
VERSION DV138221
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Euphorbia esula
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbieae; Euphorbia.
1 (bases 1 to 681)
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector). . . TAGCTTGATATCG (End Vector) (Start
EcoRI adaptor) AATCCATTTGTTGGG (End EcoRI adaptor) (Start
Tag) (Start NotI site/Vector) GCGCGCCACCGCGG. . . Base
Calling/Quality Scores: PHRED from Washington University
Center. Vector Trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG (T7)
BACKWARD: ATTAACCCCTCACTAAAG (T3)
Insert Length: 681 Std Error: 0.00
Plate: CV03119A2G06.f1 row: A column: 05
Seq primer: TAATACGACTCACTATAGG (T7)
High quality sequence stop: 681.

PCR Primers

FORWARD: TAATACGACTCACTATAGG (T7)

BACKWARD: ATTAACCCCTCACTAAAG (T3)

Insert Length: 681 Std Error: 0.00


```
QY 885 TGCAATGTTTACTCGCGGAGCAAACTTCGAAAGCTACAGAAATGGTAAATGGCAG 944
Db 824 TGGAGATGTTTATGTTGGAGTGAACGTCATCAACGATTTTGGAAATGGATCTGGCGG 765
QY 945 AGCTGATGAAATTCACACTGAATTAAGAAAGCAAGAGAGTGTAGACAAGTATTG 1004
Db 764 AATTGATGAAATCCCAAGTGTAAATGGCAAGGCAAGAAAGAGTGGGCAAGTATTGG 705
QY 1005 GTGAAATCGGAAAGTTCATGAATCAAGATTTTCATGATTTGAAATTTCTCAAGTTAGTGG 1064
Db 704 GTACTATTGAGATGTTGATGATCAATGATCGTAAATTTGAATTTCTTGAAGCTAATTA 645
QY 1065 TTAAGAAACTCTAAGATTACATCTCCGTTGCTTCGATTCGAGGAGGTGTAGAGAAA 1124
Db 644 TCAAGAAACTTTTGAGATTTTCATCTCCAGGACCTTTTCATTCCAAGAGTATGCGTCAATA 585
QY 1125 CAACAGAAATGATGGATATGAAATTTTCATCCGAACACTCGAATTG---TTGTGAATGCTT 1181
Db 584 GTTGTGAAGTTTCAACGATATGCCATAGAAGCTAACAGCAAAATTAGTAATGGTTAGTGCAT 525
QY 1182 GGGCGATAGGAGAGATCCTAATACTTGTGCGAACCTTGGAAAGTTTAAACCCAGAAAGGT 1241
Db 524 GGGCTATTGGAAAGGATCCTAATTAATGATCGAAACCTGAGAAATTCCTCTGAGAGAT 465
QY 1242 TTAAGATTTGTGCAATTGATTATAAAGGAGCAGACATTTGAACTGGTACCATTTGGTGCAG 1301
Db 464 TCCTAGATAGCTCCGTTGATTATAAGGTGCCAATTTTGAATTCATCCCATTTGGTTCTG 405
QY 1302 GAAAGAAATATGCTCTGCATTAATGAGTATTAACCAATTTGAGATGTCATTATAA 1361
Db 404 GAAGAGAAATCTGCCAGGAAATGAATCTTGGTTGGATAATGTGAGGTGGCTCTTGGCG 345
QY 1362 ATCTATTATATCATTTTAATTTGGGAACCTGGCGGATGGAATTAACCTCAAAACACTTGATA 1421
Db 344 RACTGTTATACATTTTGTGGAACCTTCCAATGGAGTAACACAGAGCTCTTGATA 285
QY 1422 TGACTGAAGCTATTGGCGGTGCTCTCAGGAAATTAAGATTTAAAGTTGATTCCTATTTC 1481
Db 284 TGGAGGAGCATTTATAGTTCCCAACTAGAAAGAGCATGATCTTATATTATTTCCCATCC 225
QY 1482 CATATC 1487
Db 224 CATATC 219
```

```
RESULT 12
LOCUS DV119846
DEFINITION DV119846 706 bp mRNA linear EST 30-SEP-2005
CV0302481C05.f2 CV03-normalized library Euphorbia esula cDNA clone
CV0302481C05.f2 5, mRNA sequence.
ACCESSION DV119846
VERSION DV119846.1 GI:76604495
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbieae; Euphorbia.
1 (bases 1 to 706)
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
```

```
(Vector)...TAAGCTTGATATCG(End Vector)(Start
EcoRI adaptor)AATTCATGTTGGG(End EcoRI adaptor)(Start
Insert)...AAAAAAAAAAAAAAAAA(End Insert)(Start Tag)TGCGT(End
Tag)(Start NotI site/Vector)GCGCGCCACCGCGG...Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector Trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
```

```
PCR Primers
FORWARD: TAATACGACTCACTATAGGG(T7)
BACKWARD: ATTAAACCTCACTAAAG(T3)
Insert Length: 706 Std Error: 0.00
Plate: CV0302481 row: C column: 05
Seq primer: TAATACGACTCACTATAGGG (T7)
High quality sequence stop: 706.
```

FEATURES

source

```
1..706
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clones="CV0302481C05.f2"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/notes="Vector: pBS II SK(+); Site 1: EcoRI (5' side of
insert); Site 2: NotI (3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse plants, and seedlings. The
library was constructed and normalized as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGT."
```

ORIGIN

```
Query Match 15.8%; Score 274.2; DB 10; Length 706;
Best Local Similarity 68.6%; Pred. No. 3.5e-59; Indels 0; Gaps 0;
Matches 375; Conservative 0; Mismatches 172;
QY 942 CAGAGCTGTGAAAAATCCAACTGAACCTAAAGAAAGCAAGAAAGTGTAGACAAGTAT 1001
Db 1 CAGAACTNNTNTACACCCAGAACATGAAGAAAGCAAGATGAGTACGACAAGTTT 60
QY 1002 TTGTGAAATCGGAAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCAAGTTAG 1061
Db 61 TTGCAGAAAAAGGAAATGTAGAAGAAATCACAACTTCACCAATTTAAATTAATTTGAAATTTAG 120
QY 1062 TGGTTAAAGAACTCTAAGATTACATCTCCGGTTGCTTCGATTCGAGGAGTGTAGAG 1121
Db 121 TCATCAAGAAACCCCTCAGATTACACCCAGCAGTCCCTTTAATCCAGAGAAATCGAGAG 180
QY 1122 AAACAACACGAATTTGATGGATATGAAATTCATCCGAACACTCGAATTTGTTGAATGCTT 1181
Db 191 AAAAAACGAAGTCAATGATATGATTTCTCCCAAAAACTAAGGCACCTTGTGAATGTTT 240
QY 1182 GGGCGATAGGAAAGAGATCCTAATACTTGGTGGAAACCTTGGAAAGTTTAAACCCAGAAAGT 1241
Db 241 GGGCAATTTCTAGGGACCCAAAAATTTGGCCCTGAAGACAGATAAAATTTATACCTGAAAGAT 300
QY 1242 TTAAGATTTGCAATTTGATTATAAGGAGCAGACATTTGAACTGTCACCATTTGGTGCAG 1301
Db 301 TTGAAAAATAGTTCAATAGATTTTAAGGGAATAATTTGGAATTCCTCGCTTTGGTTTCAG 360
QY 1302 GAAAAAGAAATATGCTCTGGCATTAATTCAGCTATTACCAATTTGGAGTATGTCAATTATAA 1361
Db 361 GAAAAAGAAATATGTCAGCATGCGTTGGGGGTAACTAAATTTGAGCTTTTCTTTCGAC 420
QY 1362 ATCTATTATATCATTTTAAATTTGGGAACCTGGCGGATGGAATTAACACTCAACACTTGATA 1421
Db 421 AGCTTCTGTATCATTTTACTGAGAACTTGCCGATGGAAAAAGACTGTGCGGAACTTGACA 480
```

QY 1422 TGACTGAAGCTATTGGCGGTCTCTCAGGAAAAAATAGATCTTAAGTTGATTCTCTATTC 1481
 |||||
 Db 481 TGGGTGAAGTAGTTGGTGGCTATTAAAAAGAAAGTGGATCTCAATTTGATTCTCTATTC 540
 |||||
 QY 1482 CATATCA 1488
 |||||
 Db 541 CATTCCA 547

RESULT 13 CX053093/c

LOCUS

DEFINITION
 UCRC09_49H06_b Ruby Orange Developing Seed cDNA Library UCRC09

VERSION
 Citrus sinensis cDNA clone UCRC09-49H06-012-1-7.b, mRNA sequence.

ACCESSION
 CX053093

KEYWORDS
 CX053093.1 GI:56537384

SOURCE
 EST.

ORGANISM
 Citrus sinensis

REFERENCE
 Citrus sinensis

AUTHORS
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

TITLE
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

JOURNAL
 Rosids; eurosids II; Sapindales; Rutaceae; Citrus.

COMMENT
 1 (bases 1 to 902)

1
 Close, T.J., Roose, M.L., Federici, C.F., Fenton, R.D., Watkins, P.,
 Morse, J., Wanamaker, S., Lyon, M., Ye, X.R., Jang, C., Quintilio, C.,
 Ikeda, J., Collin, M., Kacar, Y., Landry, B., Hubert, N., Laforest, M.,
 Landry, J. and Ligon, A.

Development of EST Resources and New Genetic Markers for California
 Citrus - Ruby Orange Developing Seed cDNA Library UCRC09

Unpublished (2004)
 Contact: Timothy J. Close

Department of Botany & Plant Sciences
 University of California

Riverside, CA 92521-0124, USA

Tel: 909-787-3318

Fax: 909-787-4437

Email: timothy.close@ucr.edu

Seq primer: T7.

Location/Qualifiers

1..902

/organism="Citrus sinensis"

/mol_type="mRNA"

/cultivar="Ruby"

/db_xref="taxon:2711"

/clone="UCRC09-49H06-012-1-7.b"

/tissue_type="Seed"

/dev_stage="20 year old trees"

/lab_host="E. coli TJC121"

/clone_lib="Ruby Orange Developing Seed cDNA Library UCRC09"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; Two trees with
 open-pollinated flowers in a mixed planting in the UC
 Riverside Citrus Variety Collection were the source of
 tissue. Developing seeds were collected by Federici (Roose
 lab) from May-July 2003. This included nine stages, based
 on size of fruit: 10-20 mm, 20-30 mm, 30-35 mm, 35-40 mm,
 40-45 mm, 45-50 mm, 50-55 mm, 55-60 mm, 60-65 mm. Tissues
 were stored in RNA Later (Ambion) until further
 processing. Fenton (Close lab) purified RNA by the phenol
 method described in J. Japanese Soc. Hort. Sci. 1996. 64
 (4):809-814, purified poly(A) mRNA using an Oligotex mRNA
 Kit (Qiagen), produced a primary cDNA library using a
 lambda ZAP XR cDNA Synthesis Kit (Stratagene), then
 mass-excised 0.48 million pfu from the primary library to
 produce a phagemid population. The library was made from
 equal portions of RNA from seed at each of the nine fruit
 stages, but seeds were not of uniform size at any stage.
 Phagemids were plated, plasmid DNA purified, cDNA clones
 archived, and DNA sequences determined bi-directionally
 using an ABI3730 at DNA Landmarks (Landry, Hubert,
 Laforest, Landry, Ligon). Chromatogram files were
 downloaded by FTP by Close, then processed by Wanamaker

FEATURES

source

(Close lab) using the HarVest pipeline
 (http://harvest.ucr.edu) to remove vector and cloning
 oligo sequences and various contaminants, and to trim to a
 high quality region. Sequences that retained a phred 17
 region of at least 100 bases were assembled, then chimeras
 were removed following manual inspection of assemblies
 (Close, Roose, Federici, Wanamaker, Lyon, Ye, Jang,
 Collin, Kacar, Ikeda, Quintilio). Sequences that survived
 all removal steps were submitted to GenBank."

ORIGIN

Query Match 15.8%; Score 273.2; DB 8; Length 902;
 Best Local Similarity 62.5%; Pred. No. 6.7e-59;
 Matches 444; Conservative 0; Mismatches 263; Indels 3; Gaps 1;
 QY 782 AAGCCTTTTGAAGCGGATAACTTAATGGATGTTCTATTGNAATCTTCAAAAAAATGGAAC 841
 |||||
 Db 826 AAGACTGGTGATGATGAAGATCTGGTTGATGTTCTGTTGAAAAATCAAGGGCTCGGAAAT 767
 |||||
 QY 842 GTTCCAGTCCCAAGTGACAAACGAAGCATCAAAAGCATCCGTTTTCAGAAATGTTTACTGCC 901
 |||||
 Db 766 TATGGTTCTTCTTGACTACACACATCAAGCAGTAATCTCTGCATTTTGGTGCT 707
 |||||
 QY 902 GGGAGCGAAACAACTTCGAAAGCTACAGAATGGTAATGGCAGAGCTGATGAAAAATCCA 961
 |||||
 Db 706 GGGAGTGAACACATCTGCCACCACCTGTGGATTGGCAATGTGTGAATGATAAAAAATCCA 647
 |||||
 QY 962 ACTGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTAAAGAACTCTTAAGA 1081
 |||||
 Db 646 AGAGTTTAAAAAAGCAGAAAGCTGAGCGAGGAAGTGTGTTAATAGAGGGGGAAGTC 587
 |||||
 QY 1022 GATGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTAAAGAACTCTTAAGA 1081
 |||||
 Db 586 GAGCAACCGGGCATTAAGAAATGATGTACATGAAGCTAGTTGTTAAAGAGACATTTGAGA 527
 |||||
 QY 1082 TTACATCTCTCGG---TTGTCCTTGATTCCGAGGGAGTGTAGAGAAACACACGAATGTAT 1138
 |||||
 Db 526 CTACATCTCTCTGGTCTTTGTTATTGCAAGAGAAATGTGGAGAGCGATGTAAGATTAA 467
 |||||
 QY 1139 GGATATGAATTCATCCGACACTCGAATTTGTTGAATGCTTGGCGGATAGAGAGAT 1198
 |||||
 Db 466 GGATTTGACATACCTGTGAAAGCTAAAGTCATTTGTTAATGTCATGGTGTGATGGAGAGAT 407
 |||||
 QY 1199 CCTAATACTTGTTCGGAACCTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATT 1258
 |||||
 Db 406 CCAAGTACTGACCGAACCTCAGAGCTTTATTCAGAGAGGTTCTCTCATTTATTCTGTT 347
 |||||
 QY 1259 GATTATAAGGGACGACATTTTGAACCTGGTACCATTGTTGGTCAGGAAAAAGATATGCTCT 1318
 |||||
 Db 346 GACTACAAGGGGACTAAATTTTTCGGTACATCCCATTTGGTGTGTTAGGAGGATCTGCCCT 287
 |||||
 QY 1319 GGATTTACTTCAAGCTATTACCAATTTGGAGTATGTCATTTATAATCTATTATATCATTTT 1378
 |||||
 Db 286 GGAACGTCTATTTGGTCTGGCCAGCATGGAGCTTACACTCGCATGTTTATTATACCATTC 227
 |||||
 QY 1379 AATTGGGAACCTGCGCATGGAATTAACACCTCAAACTTTGATATGATGATGATGATTTGGC 1438
 |||||
 Db 226 GAATGGAACCTACCAATTTGGAATGGAATGAGCATGAGGATTTGGACATGACTGAGGCTTTGGC 167
 |||||
 QY 1439 GGTGCTCTCAGGAAAAAATAGATCTTAAGTTGATTTCCTTATTCATATCA 1488
 |||||
 Db 166 GCAACAGTAAAGAAAGAAACAAAGATTTGTGTGATGATTTCTTATTCACACCA 117
 |||||

RESULT 14

DV128282

LOCUS

DEFINITION

CV03050B1D03.f1 CV03-normalized library Euphorbia esula cDNA clone

VERSION

DV128282

KEYWORDS

SOURCE

DV128282 674 bp mRNA linear EST 03-OCT-2005
 CV03050B1D03.f1 CV03-normalized library Euphorbia esula cDNA clone
 CV03050B1D03.f1 5, mRNA sequence.
 DV128282
 DV128282.1 GI:76837284
 EST.
 Euphorbia esula (leafy spurge)

ORGANISM Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae; Euphorbieae; Euphorbia.

REFERENCE
AUTHORS Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A., Kim, W.R. and Mikel, M.
TITLE Direct submission, Anderson, J.V. 2005
JOURNAL Unpublished (2005)
COMMENT Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND 58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
ECORI adaptor/TAATCCATTGTGTGG (End EcorI adaptor) (Start Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TCGGT (End Tag) (Start NotI site/Vector) GCGCGCCGCCCGCGG . . . Base Calling/Quality_Scores: PHRED from Washington University Genome Center. Vector Trimming: Cross_match from Washington University Genome Center PHRAP suite. Low quality bases (Phred score < 20) were trimmed from both ends of the sequence by an in-house script. This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG (T7)
BACKWARD: ATTAACCTCTACTTAAG (T3)
Insert Length: 674 Std Error: 0.00
Plate: CV03050B1 row: D column: 03
Seq primer: TAATACGACTCACTATAGG (T7)
High quality sequence stop: 674.

FEATURES source

FEATURES
1..674
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03050B1003.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcorI (5' side of insert); Site 2: NotI (3' side of insert); Tissue: Seasonal crown buds from outdoor plants, whole plant tissue (leaf, stem, flowers, seeds, shoots, galls, root and crown buds) from outdoor plants, cold acclimated tissues (leaf, stem, and crown buds), crown and root buds from decapitated greenhouse p ants, and seedlings. The library was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. An identifying tag was added at the 3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAAATCGGT."

ORIGIN
Query Match 15.7%; Score 271.8; DB 10; Length 674;
Best Local Similarity 64.0%; Pred. No. 1.4e-58;
Matches 427; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY 428 TTTCTGAGTCCAAAAGTTCAATCCCTCCAGTTAATCCGAGAGAGAAATGGAGAT 487
DB 1 TTTCTGAGTCTGAAACGTTCAATCTTTTCAGATCTCTCAGGGAAGAAACGTTTCAGGT 60

QY 488 GCATCACATTCCTCGGTCGAAAGCCGGATCTCCGGTCAATATTACAAAGATCATTTAT 547
DB 61 TTTATCAGATTCCTCGGTACAAAGCAAAATTCGTTGTAATCTGACGAAATCCGTCGGT 120

QY 548 GCATTAATAATTTCCGATCATGATAGAACATCCGTTGGTAAATGTAAGCAAAAAGAAAGA 607
DB 121 GATTTCGACAGTTCAATATGCTTATTTAACTTATGGAATAATGTGATGAAAAATTTGTTG 180

QY 608 TTGCTGAGTGTGGCCGATGTCAGTCAATGAGGAGGACGAGTTTGGGACCCGACAGCT 667

Db 181 GCTATGTTGGAAAAAGTGAAGCAAGCAGTGTATAGACGACGAGTAGTGTA---CGGATCTG 237
QY 668 TTTCCGACGTGGAAATTTACTTTCTACTATATCTGAGCTGATCAAAACCCAGCGCTTTG 727
DB 238 TTTCTCTCGTTGAAATTCATTTCAGTATATCAATGTTGAGAAAGTCAAGAATGCGCAAGGTA 297

QY 728 CATCAGGAGATTGAGTATATCTTGAAGAGATTCTTAATGAACACAAAGCCCAATAGCCT 787
DB 298 CAAAAGGAAATGGATAAAATGCTTTGAAGATATTATTAAGAAACACAAAGTTTCAATATAAG 357

QY 788 TTTGAAGCGGATAACTTAATGATGTTCTATTGATCTTCAAAAAAATGGAACCTTCCA 847
DB 358 TTTGAGGATAACAATCTTTTACAGTCTTTTGAATCAACAGCAAAATGGAGATCTTGA 417

QY 848 GTGCCAGTGCACAAACGAAAGCATCAAGCATCCGTTTTTGCAAAATGTTTACTGCCCGGAGC 907
DB 418 CTTCATTTGACAAATGAAGTCATCAAGCCCAACATTATGGAATAATTTTTTGGTGGGAGC 477

QY 908 GAAACAATTCGAAAGCTACAGATGGTATGCGAGAGCTGATGAAAAATCCAATGAA 967
DB 478 CATACATCTTCTAAATTTGTGGAATGGCAATGTCAGAGTTGATGAAGAACCCAGAAATCA 537

QY 968 CTAAGAAAAGCACAAAGAAAGTGTAGACAAGTATTGTGTAATGGGAAAAGTTGATGAA 1027
DB 538 ATGACAAAGCACAAAGAGAGTGAACAATCTTTGTGAAAAGGGAAGTTCGAGGA 597

QY 1028 TCAAGATTTTCATGATTTGAAATTTCTCAAGTTAGTGGTTAAAGAAATCTTAAGATTACAT 1087
DB 598 TCAATATGCAAGAACTGAAATACCTCAAGTCAATTTCAAAAGAAATCTCAGATTGCAC 657

QY 1088 CTTCCGG 1094
DB 658 CTTCCGG 664

RESULT 15

LOCUS CX636845
DEFINITION UCRPT02_22G04 g Poncirus trifoliata Roots with Iron Deficiency - UCRPT02-UCR1 Poncirus trifoliata cDNA clone UCRPT02-22G04-M7-1-5.g, mRNA sequence.

ACCESSION CX636845
VERSION CX636845.1 GI:57871674
KEYWORDS EST.
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Sapindales; Rutaceae; Poncirus.

REFERENCE 1 (bases 1 to 819)
AUTHORS Close, T.J., Roose, M.L., Parker, D.R., Federici, C.F., Mandal, J., Fenton, R.D., Wanamaker, S., Landry, B., Hubert, N., Laforest, M., Landry, J. and Ligonde, A.
TITLE Development of EST Resources and New Genetic Markers for California Citrus - Poncirus trifoliata Roots with Iron Deficiency - UCRPT02-UCR1

JOURNAL Unpublished (2005)
COMMENT Contact: Timothy J. Close
Department of Botany & Plant Sciences
University of California
Riverside, CA 92521-0124, USA
Tel: 909-787-3318
Fax: 909-787-4437
Email: timothy.close@ucr.edu
Seq primer: T3.
Location/Qualifiers
1..819
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy"
/db_xref="taxon:37690"
/clone="UCRPT02-22G04-M7-1-5.g"
/tissue_type="root"

/dev_stage="seedling"
/lab_host="E. coli TJU121"
/clone_lib="Poncirus trifoliata Roots with Iron Deficiency - UCRP02-UCR1"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plant materials were prepared by Federici (Roose lab) with advice from Parker. Seedlings 138 days after sowing were about 20-23 cm in height, bearing 8-12 leaves. On May 26, 2004, plants were washed free of soil with a stream of water, then placed upright with the roots submerged in two tanks of nutrient solution. The solution was sufficient in all major and minor nutrients and buffered with MES at about pH 6.95. The plants were maintained in this until June 14, 2004, when the solution was changed to one with only 20 micromolar iron, chelated with EDTA to induce iron deficiency. The pH was maintained at 6.99 by sodium carbonate/CO2 buffering. This solution was replaced on July 6, 2004. Roots from three plants were sampled on June 16, June 21, July 1 and July 28, 2004. Roots were collected by removing the plant from the nutrient solution, blotting off excess moisture with a paper towel, then cutting off the top of the plants. Three plants were pooled in one aluminum foil packet, and frozen between two sheets of dry ice. The time between removal from solution and freezing on dry ice did not exceed one minute. The frozen tissue was stored in these foil packs at -80°C. This sampling strategy did not correspond to initiation of stress in the plant, but only to when the plants entered the low iron solution. The actual stress was not initiated until the internal iron reserve was depleted. By the July 1 sampling date, slight iron deficiency symptoms were apparent. By the final sampling date, clear iron deficiency symptoms were present. By the July 1 sampling date, the roots showed growth of *Thielaviopsis brevicola*, a fungus. Other than what came away when the roots were blotted with paper towels, no effort was made to remove the fungus because it is not just a surface contaminant, but grows within the roots as well. Mandal and Fenton (Close lab) purified RNA using TRIzol, poly(A) mRNA using an Oligotex mRNA Kit (Qiagen), produced a primary cDNA library using a lambda ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised 0.5 million pfu from the primary library to produce a phagemid population. The library was made from equal portions of RNA from each of the four collection dates. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at DNA Landmarks (Landry, Hubert, Laforest, Landry, Ligonde). Chromatogram files were downloaded by FTP by Close, then processed by Wananaker (Close lab) using the HarVest pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were assembled, then chimeras were removed following manual inspection of assemblies (Close, Roose, Wananaker). Sequences that survived all removal steps were submitted to GenBank."

ORIGIN

Query Match 15.6%; Score 270; DB 9; Length 819;
Best Local Similarity 60.7%; Pred. No. 4.4e-58;
Matches 460; Conservative 0; Mismatches 295; Indels 3; Gaps 1;
QY 791 GAAGCGGATACTTAATGATGTTCTATTGAACTCTTCAAAAAATGGAAACGTTCCAGTG 850
DB 30 GAAGTTGAAGATCTTGTGATGTTCTGTGAAAGCTCAAGAGGATGTAATCTTGAATTT 89
QY 851 CCAGTGACAAACGAAGCATCAAGCATCCGTTTTCGAATGTTTACTCGCGGAGCGNA 910
DB 90 CCCCTAACTTACCAAAAAACGTCAAAGCTGTTGTCATGGATGTTTTTAGTGCTGGCCAG 149
QY 911 ACAACTTCGAAAGCTACAGAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACTA 970

Db 150 ACATCAGCAACAACCTGTAGATTGGGCTATGCTCTTGTGATGAGTATCCAAAAATAATG 209
QY 971 AGAAAAAGCAAGAAGAAGTAGTAGACAAGTATTTGGTGAATAATGGGAAAAAGTTGATGAATCA 1030
Db 210 AAAAAAGGCACAGATGAGGTAAAGGAAAGTCTTTTAATAGAAAGGGGAAAGTCGATGAAGCG 269
QY 1031 AGATTTTCATGATTTCGAAATTCCTCAAGTTTAGTGGTTTAAAGAAACTCTTAAGATTACATCCT 1090
Db 270 GGCCTTGAGGAATGAAGTACTTTGAAGCTAGTTTAAAGAGACTCTGAGAATACAGCCT 329
QY 1091 CCGGTT---GTCTTGATTCCGAGGGAGTGTAGAGAAAAACACAGAAATTTGATGATGAATGAA 1147
Db 330 TCTGCTCCAAATGTTATTTCCACGAGAGTGTGCAGAGACGTGTAATATTAAATGATACGAC 389
QY 1148 ATTCATCCGAACACTCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCCTTAATACT 1207
Db 390 ATACAGCCAAAGCAAGTACTTGTTAATGTATGGCGGATGACAAGAGATTCCCAATAT 449
QY 1208 TGGTCGGAACTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAA 1267
Db 450 TGGACTGAACCTGAGAGCTTTATGCGAGAGAGTTCCTTGATTCCTATAGACTTTAAG 509
QY 1268 GGGACGACATTTGAACACTGGTACCATTTGGTGCAGGAAAAAGAAATATGTCCTGGCATTACT 1327
Db 510 GGTCTCTAATTTTGAATTTCTCCCATTTGGTCTGGTAGGAGGATGTGCCCTGGAAATGACA 569
QY 1328 TCAGCTNTATCAAAATTTGGAGTATGTCATTATATAATCTATTATATCATTTTAATTTGGAA 1387
Db 570 TATGCAATGCCAACATTTGAGCTTCCACTTGCATTTGTTTGTACCATTTTCGATTGGAAAG 629
QY 1388 CTGGCCGATGGAATTCACACTCAACACTTTGATATGACTGAAGCTATTGGCGGTGCTCTC 1447
Db 630 CTACCAATGGAAATGAGAAATGAGGACTTAGACATGACAGAGGATTTGGTGTGACCGTC 689
QY 1448 AGGAAAAAATAGATCTTTAAGTTGATTCCTATTTCCATATCAAGTTAGCTTAGGCTCAAAAT 1507
Db 690 AAAAGAAAGCAAAATCTGTGCTGATTCCCACTCTCTTATATATCTCCAGCTGTTCATATA 749
QY 1508 ATTTCTTGATTACATAGGAGGTTGAAATATATATAAT 1545
Db 750 GCTCTAGAAGGAAAAAACTACAGTATAATTTTGAATAAT 787

Search completed: May 31, 2006, 14:15:58
Job time : 8723 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 08:41:59 ; Search time 1147 Seconds
(without alignments)
10534.356 Million cell updates/sec

Title: US-10-759-813-1

Perfect score: 1733

Sequence: 1 gcataaagggaatggagc.....tcttttcaaatccgaaaaa 1733

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_8:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005s:*
- 15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733	100.0	1733	6	Abk11137 cDNA enco
2	410	23.7	1538	12	Adn10568 Nicotiana
3	410	23.7	1538	14	Adz65256 Tobacco f
4	410	23.7	1538	14	Adz65803 Tobacco c
5	410	23.7	1538	14	Aee06775 Tobacco c
6	410	23.7	1538	15	Aef97168 Tobacco p
7	400.4	23.1	1673	12	Adn10646 Nicotiana
8	400.4	23.1	1673	12	Adn10624 Nicotiana
9	400.4	23.1	1673	12	Adn10626 Nicotiana
10	400.4	23.1	1673	14	Adz65334 Tobacco f
11	400.4	23.1	1673	14	Adz65312 Tobacco f
12	400.4	23.1	1673	14	Adz65314 Tobacco f
13	400.4	23.1	1673	14	Adz65861 Tobacco c
14	400.4	23.1	1673	14	Adz65859 Tobacco c
15	400.4	23.1	1673	14	Adz65881 Tobacco c
16	400.4	23.1	1673	14	Aee06853 Tobacco p
17	400.4	23.1	1673	14	Aee06831 Tobacco p
18	400.4	23.1	1673	14	Aee06833 Tobacco p

19	400.4	23.1	1673	15	Aef97224	Tobacco c
20	400.4	23.1	1673	15	Aef97226	Tobacco c
21	400.4	23.1	1673	15	Aef97246	Tobacco c
22	399.4	23.0	1781	4	Adn15839	Pepper cy
23	399.4	23.0	1781	7	Ado58604	Pepper cy
24	398.8	23.0	1673	12	Adn10628	Nicotiana
25	398.8	23.0	1673	14	Adz65316	Tobacco f
26	398.8	23.0	1673	14	Adz65863	Tobacco c
27	398.8	23.0	1673	14	Aee06835	Tobacco p
28	398.8	23.0	1673	15	Aef97228	Tobacco c
29	393.2	22.7	1610	12	Adn10676	Nicotiana
30	393.2	22.7	1610	12	Adn10612	Nicotiana
31	393.2	22.7	1610	14	Adz65300	Tobacco f
32	393.2	22.7	1610	14	Adz65364	Tobacco f
33	393.2	22.7	1610	14	Adz65911	Tobacco c
34	393.2	22.7	1610	14	Adz65847	Tobacco c
35	393.2	22.7	1610	14	Aee06819	Tobacco p
36	393.2	22.7	1610	14	Aee06883	Tobacco p
37	393.2	22.7	1610	15	Aef97276	Tobacco c
38	393.2	22.7	1610	15	Aef97212	Tobacco c
39	388.8	22.4	1566	12	Adn10606	Nicotiana
40	388.8	22.4	1566	14	Adz65294	Tobacco f
41	388.8	22.4	1566	14	Adz65841	Tobacco c
42	388.8	22.4	1566	14	Aee06813	Tobacco p
43	388.8	22.4	1566	15	Aef97206	Tobacco c
44	388.8	22.4	1576	12	Adn10564	Nicotiana
45	388.8	22.4	1576	14	Adz65252	Tobacco f

ALIGNMENTS

RESULT 1

ABK11137
ID ABK11137 standard; cDNA; 1733 BP.

XX ABK11137;

XX 05-JUN-2002 (first entry)

XX cDNA encoding Euphorbia lagascae cytochrome P450 enzyme.

XX Cytochrome P450 enzyme; synthesis of delta 12-epoxy fatty acid; epoxide;
KW vernolic acid; modified fatty acid; oil; commercial crop; plasticiser;
KW crosslinking coating application; setting printing ink; transgenic;
XX plant; EST; clone cellc_pk002_i4; gene; ss; expressed sequence tag.

XX Euphorbia lagascae.

XX Key Location/Qualifiers

XX CDS 14..1516
XX FT /*tag= a
XX FT /product= "Cytochrome P450 enzyme"

XX WO200208269-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US022790.

XX 21-JUL-2000; 2000US-0219833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB;

XX WPI; 2002-257331/30.

XX P-PSDB; AAU77470.

XX Novel isolated polynucleotide encoding plant cytochrome P450 enzyme
PT associated with synthesis of delta12-epoxy fatty acids, useful for
PT creating transgenic plants with higher or lower level expression of the
PT enzyme.

XX Claim 8; Page 49; 53pp; English.

XX The present invention relates to the isolation of the polynucleotide and

XX polypeptide sequences for a plant cytochrome P450 enzyme associated with

CC the synthesis of delta 12-epoxy fatty acids from Euphorbia lagascae. The

CC polynucleotide sequence of the invention can be used to produce epoxide

CC containing fatty acids such as vernolic acid. The sequences of the

CC invention can be used to manipulate modified fatty acids to produce oils

CC in commercial crops. They can also be used to produce plasticisers, for

CC crosslinking coating applications, and setting printing inks. The

CC polynucleotide can also be used for creating transgenic plants in which

CC the enzyme is present at higher or lower levels than normal, in cell

CC types or in developmental stages in which they are not normally found.

CC The present sequence encoding the E. lagascae cytochrome P450 enzyme of

CC the invention is isolated from the expressed sequence tag (EST) clone

XX eellic.pk002.14

SQ Sequence 1733 BP; 556 A; 282 C; 376 G; 519 T; 0 U; 0 Other;

Query Match 100.0%; Score 1733; DB 6; Length 1733;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATAAAGGAAATGGAGCAGAAAATCTCTCTTTCCGAGCATTTTAAATAGTTTTCT 60

DB 1 GCATAAAGGAAATGGAGCAGAAAATCTCTCTTTCCGAGCATTTTAAATAGTTTTCT 60

QY 61 GCTTGTGTTTAACTTCTAGTAGTATCATGAGGTTGTGGAAGAACAGAAATCCACCTCCAGG 120

DB 61 GCTTGTGTTTAACTTCTAGTAGTATCATGAGGTTGTGGAAGAACAGAAATCCACCTCCAGG 120

QY 121 GCCATGGAGTTTCTCTATCATAGGTAATCTTCTCATTTTATTTACTCCTCTGATCTAGG 180

DB 121 GCCATGGAGTTTCTCTATCATAGGTAATCTTCTCATTTTATTTACTCCTCTGATCTAGG 180

QY 181 CCATGAAAGTTTTAGAGCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAATGG 240

DB 181 CCATGAAAGTTTTAGAGCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAATGG 240

QY 241 CCAAGTTTCAGCTGTTGTTCATTTCTCAGCTCAAGCAGCCAAAGAGGTTATGAAAACCTCA 300

DB 241 CCAAGTTTCAGCTGTTGTTCATTTCTCAGCTCAAGCAGCCAAAGAGGTTATGAAAACCTCA 300

QY 301 GGCTGATGCTTTCCGCCAACGCCCTATCGTCTTGGACGCACAGATTGTTTTATAATCG 360

DB 301 GGCTGATGCTTTCCGCCAACGCCCTATCGTCTTGGACGCACAGATTGTTTTATAATCG 360

QY 361 GAAAGATGCTTTGTTGCTTCATATGGAGATCACTGGAGGAGATGAAGAAAATTTGGAT 420

DB 361 GAAAGATGCTTTGTTGCTTCATATGGAGATCACTGGAGGAGATGAAGAAAATTTGGAT 420

QY 421 ACTTCGAATTTCTGAGTGCCAAAAGTTTCAATCTCCAGGTTAATCCGAGAGAGAAAT 480

DB 421 ACTTCGAATTTCTGAGTGCCAAAAGTTTCAATCTCCAGGTTAATCCGAGAGAGAAAT 480

QY 481 GGAGGATGCCATCACATTTCTCCGTTGAAAGCCGATCTCCGTTCAATATTACAAAGAT 540

DB 481 GGAGGATGCCATCACATTTCTCCGTTGAAAGCCGATCTCCGTTCAATATTACAAAGAT 540

QY 541 CATTTATGGCATTAATTTTCATCATGATGAAGACATCCGTTGGTAAATTTGAACAAA 600

DB 541 CATTTATGGCATTAATTTTCATCATGATGAAGACATCCGTTGGTAAATTTGAACAAA 600

QY 601 AGAAGAGTTGCTGAGTTTGGCCGATGAGTCAATGAGGAGCAGCGAGTTTGGCACCGC 660

DB 601 AGAAGAGTTGCTGAGTTTGGCCGATGAGTCAATGAGGAGCAGCGAGTTTGGCACCGC 660

QY 661 AGACGCTTTCCGAGCGTGAATTTACTTCACTATATCATTTGAGCTGAGTCAAAACCCAG 720

DB 661 AGACGCTTTCCGAGCGTGAATTTACTTCACTATATCATTTGAGCTGAGTCAAAACCCAG 720

QY 721 GGGTTGCATCAGGAGTTGACGATATATCTTGAAGAGATTTCTTAATGACACAAAGCCAA 780

RESULT 2
ADN10568
ID ADN10568 standard; cdNA; 1538 BP.
XX
AC ADN10568;

DB 721 GCGTTTCATCAGGAGATTGCGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA 780

QY 781 TAAGCCTTTTGAAGCGGATACTTAATGGAATGTTCTTATGGAATCTTCAAAAATGGA 840

DB 781 TAAGCCTTTTGAAGCGGATACTTAATGGAATGTTCTTATGGAATCTTCAAAAATGGA 840

QY 841 CGTTCCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTGC 900

DB 841 CGTTCCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTGC 900

QY 901 CGGAGCGAAACAACTTCGAAAGCTACAGATGGTAAATGGCAGAGCTGATGAAAATCC 960

DB 901 CGGAGCGAAACAACTTCGAAAGCTACAGATGGTAAATGGCAGAGCTGATGAAAATCC 960

QY 961 AACTGAACTAAGAAAAGCAAGAAAGTAGACAAGTATTGGTGAATGGGAAAAGT 1020

DB 961 AACTGAACTAAGAAAAGCAAGAAAGTAGACAAGTATTGGTGAATGGGAAAAGT 1020

QY 1021 TGATGAATCAAGATTTTCATGATTTTGAATTTCTTCAAGTTAGTGGTTAAGAAAATCTTAAG 1080

DB 1021 TGATGAATCAAGATTTTCATGATTTTGAATTTCTTCAAGTTAGTGGTTAAGAAAATCTTAAG 1080

QY 1081 ATTACATCTCCGTTGCTTCTGATTCCGAGGAGGTAGAGAAAACAAACAGAAATGATGG 1140

DB 1081 ATTACATCTCCGTTGCTTCTGATTCCGAGGAGGTAGAGAAAACAAACAGAAATGATGG 1140

QY 1141 ATATGAAATTCATCCGAAACACTCGAAATTTGTGTAATCTTGGGCGATAGGAAGATCC 1200

DB 1141 ATATGAAATTCATCCGAAACACTCGAAATTTGTGTAATCTTGGGCGATAGGAAGATCC 1200

QY 1201 TAATACCTTGGTCCGAACTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTCGAATTGA 1260

DB 1201 TAATACCTTGGTCCGAACTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTCGAATTGA 1260

QY 1261 TTATAAGGGAGCAGCATTTGAACTGGTACCATTTGGTGCAGAAAAGAAATATGTCCTGG 1320

DB 1261 TTATAAGGGAGCAGCATTTGAACTGGTACCATTTGGTGCAGAAAAGAAATATGTCCTGG 1320

QY 1321 CATTACTTCAGCTATTACCAATTTGGAGTATGTCATTAATTAATCTATTATATCATTTTAA 1380

DB 1321 CATTACTTCAGCTATTACCAATTTGGAGTATGTCATTAATTAATCTATTATATCATTTTAA 1380

QY 1381 TTGGAACTGCCGATGGAATTTACCTCAACACCTTGATGACTGAAGCTATTGGCGG 1440

DB 1381 TTGGAACTGCCGATGGAATTTACCTCAACACCTTGATGACTGAAGCTATTGGCGG 1440

QY 1441 TGCTCTCAGGAAAAAATAGATCTTAAAGTTGATTCTTATTCATATCAAGTTAGCTTAGG 1500

DB 1441 TGCTCTCAGGAAAAAATAGATCTTAAAGTTGATTCTTATTCATATCAAGTTAGCTTAGG 1500

QY 1501 CTCAAATATTTCTTGATTAACATAGAGGGTTGAAATATATATAATAATAATAATAAAG 1560

DB 1501 CTCAAATATTTCTTGATTAACATAGAGGGTTGAAATATATATAATAATAATAATAAAG 1560

QY 1561 ATGTTCTTAATGTTTGGGTTAGTTAATAGGTTTCCACCGATCATATAAGTAGCCT 1620

DB 1561 ATGTTCTTAATGTTTGGGTTAGTTAATAGGTTTCCACCGATCATATAAGTAGCCT 1620

QY 1621 TCTTTGATGGATGGGTTAGATTATAATGAGTTGCTGGGTTGGATTTTAGATGGGTTAAAT 1680

DB 1621 TCTTTGATGGATGGGTTAGATTATAATGAGTTGCTGGGTTGGATTTTAGATGGGTTAAAT 1680

QY 1681 GATTGGATGGAATAATAATAAATGAAATGTTTCTTTTCAATCCGAAAAA 1733

DB 1681 GATTGGATGGAATAATAATAAATGAAATGTTTCTTTTCAATCCGAAAAA 1733

[illegible]

Db 1143 CCAAGAGAAATGTAGGAAGAGACAATAATATAAACGGCTACACTATTCCTGTTAAAGACCAAA 1202
Qy 1166 ATTGTTGTCAATGCTTTGGCGCATAGGAAGAGATCCCTAATACCTTGTGCGAACCTCGAAAG 1225
Db 1203 GTCATGGTTAATGTTTGGCGATTGGAGAGATCCAAATATTTGGATGATGCAGAACT 1262
Qy 1226 TTAAACCAAGAAAGTTTAAAGATTGTGCAATTGATTAATAAGGAGCAGACATTTGAATG 1285
Db 1263 TTAAAGCAGAGAGATTGAGCAGTGTCTAAGGATTTTGTGGTAATAATTTGAATAT 1322
Qy 1286 GTACCAATTTGGTCAGGAAGAAAGATATGCTCTGCATTAATCTTACGCTATTACCAATTTG 1345
Db 1323 CTTCCATTGGTGGTGAAGGAGGATTTGTCAGGGAATTCGTTGGTTAGCTAATGCT 1382
Qy 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACTGCCGATGGAAATTACA 1405
Db 1383 TATTTGGCCATTGGCTCAATTACTTTTATCATTGATTTGGAACTCCCCACTGGATCAA 1442
Qy 1406 CTTCAACACATTTGATATGACTGAAGCTATTTGGCGTGTCTCAGGAAATAAATAGATCTT 1465
Db 1443 CCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAACCTGCCGCTAGAAAGAGTGACCTT 1502
Qy 1466 AGTTTGATTCCTATTCCATATCAA 1489
Db 1503 TACTTGGTTCGACTCCTTATCAA 1526

RESULT 5

ID ABE06775
AC ABE06775 standard; cDNA; 1538 BP.
XX ABE06775;
DT 26-JAN-2006 (first entry)
XX
DE Tobacco p450 cDNA SEQ ID NO: 162.
XX
KW Plant breeding; plant; ss; coding sequence; cytochrome p450;
KW transgenic plant; senescence.
XX
OS Nicotiana tabacum.
XX
PN WO2005111217-A2.
XX
PD 24-NOV-2005.
XX
PF 27-APR-2005; 2005WO-US014803.
XX
PR 29-APR-2004; 2004US-0566235P.
PR 03-SEP-2004; 2004US-00934944.
PR 03-SEP-2004; 2004US-0607357P.
PR 17-SEP-2004; 2004US-00943507.
PR 15-OCT-2004; 2004WO-US034065.
PR 15-OCT-2004; 2004WO-US034218.
PR 25-JAN-2005; 2005US-0646764P.
PR 24-MAR-2005; 2005US-0665097P.
PR 24-MAR-2005; 2005US-0665451P.
XX
(USSN-) US SMOKELESS TOBACCO CO.
XX
XX Xu D, Nielsen MT;
PI
XX
DR WPI; 2005-786788/80.
DR P-PSDB; ABE06776.
XX

PT Producing a tobacco plant having decreased expression of a nicotine
PT demethylase gene comprises crossing a first tobacco plant with a second
PT tobacco plant and germinating the collected seed of an F1 progeny plant.

XX Claim 32; SEQ ID NO 162; 641pp; English.

XX The invention relates to a breeding method for producing a tobacco plant
CC with reduced expression of a nicotine demethylase gene comprises crossing

CC a first tobacco plant with variant nicotine demethylase gene expression
CC with a second tobacco plant with at least one phenotypic trait to produce
CC an F1 progeny plant, the seed of which is collected and germinated to
CC produce a tobacco plant having reduced expression of a nicotine
CC demethylase gene. Also included are breeding a nicotine demethylase
CC deficiency trait into a tobacco plant, producing a tobacco seed,
CC developing a tobacco plant in a tobacco breeding program, a tissue
CC culture of regenerable tobacco cells obtained from the tobacco plant of
CC the invention, producing a tobacco product, a breeding method for
CC producing a tobacco plant having a modified attribute, a method of
CC breeding an attribute into a tobacco plant, a tobacco plant or its
CC components produced by the method of breeding a nicotine demethylase
CC deficiency trait into a tobacco plant, producing tobacco seed, producing
CC a tobacco plant having a modified attribute or developing a tobacco plant
CC in a tobacco breeding program, an isolated genetic marker comprising a
CC nucleic acid sequence that is substantially identical to a nucleic acid
CC sequence given in the specification (the nucleic acids comprise isolated
CC cytochrome p450 cDNAs), an expression vector comprising the nucleic acid
CC sequence, a plant or plant component comprising the isolated nucleic acid
CC sequence, a plant produced from a germinated seed of the plant, reducing
CC the expression or enzymatic activity of a constitutive, or an ethylene
CC induced or senescence induced tobacco polypeptide in a plant cell, and
CC increasing the expression or enzymatic activity of a constitutive, or an
CC ethylene or senescence induced tobacco polypeptide in a plant cell. The
CC phenotypic trait comprises disease resistance, high yield, high grade
CC index, curability, curing quality, mechanical harvestability, holding
CC ability, leaf quality, height, maturation, stalk size, or leaf number per
CC plant. The breeding method for producing a tobacco plant having decreased
CC expression of a nicotine demethylase gene is useful developing desirable
CC (non-genetically engineered) germplasm. The plant is useful in producing
CC (smokeless) tobacco products. The tobacco product is a moist or dry
CC snuff, a chewing tobacco, a cigarette product, a cigar product, a
CC cigarillo, a pipe tobacco, or bidis. The p450 cDNAs were isolated using
CC degenerate PCR primers designed against cytochrome p450 motifs. The
CC present sequence is a cytochrome p450 cDNA of the invention.
XX
XX Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Query Match 23.7%; Score 410; DB 14; Length 1538;

Best Local Similarity 58.1%; Pred. No. 1.5e-93;

Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CCTCAGGCGCATGGAAGTTTCCTATCATAGTAATCTTCCTCATTTATTACTCACTTCT 172
Db 129 CCACGAGTCCATGGAAACTACCAATACCTAGAGATGCTTCATATGTT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAATTTATGGACCTGTTATAGTCTT 232
Db 183 GGACTACCAACCATGTCCTTAGAGATTTAGCCAAAAAATATGGACCGCTTATGCACCTT 242
Qy 233 CAATTTGCCCAAGTTTCAGCTGTTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 243 CAATTTAGGTGAAGTTTCTGTCAGTTGTGGTTACTTCTCTCGATATGCGAAAAAGACTCTA 302
Qy 293 AAAACTCAGGCTGATGCTTCGCCCAACGCTTCGTTCTTGGAGCCACAGATTTGTGTTT 352
Db 303 AAAACTCATGACATCGCTTTCGCTCTAGGCTAGCCCTTTTGGCCCCCGAGATTTGCTGT 362
Qy 353 TATAATCGGAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGCGAGATGAAGAAA 412
Db 363 TACAATAGTCTGATCTTGGCTTTTGGCCCTATGCGGATTTTGGAGACAAATGCGTAAA 422
Qy 413 ATTTGGATACATTGAATTTCTGAGTGCCAAAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
Db 423 ATATGTGCTTTGGAAGTGTCTCAGTSCCAAGAAATGTTCCGACATATAGCTCTATTATGGCGC 482
Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTCGAAAGCCGGATCTCCCGTCAATATT 532
Db 483 GATGAAGTCTTCGCTCTCTTAATTTATCCGGTCATCTTCTGTTGAGGCTCTTAAATATT 542
Qy 533 ACAAGATCATTTATGGCATTATAATTTGATCATGATGAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602

terpenoids, lipids, cyanogenic glycosides and glucosinolates, with effects on the flavor or aroma of plant products. They may also affect herbicide tolerance, resistance to disease or insects, quality factors related to undesirable constituents, structural traits, fiber content, leaf yield, ripening, leaf curing or storage properties. Tobacco plants of the invention in which the expression of cytochrome P450 genes is altered may have desirable traits such as altered levels of nicotine or N'-nitrosonicotine. Such plants can be used in the production of tobacco products such as moist or dry snuff, chewing tobacco, cigarettes, cigars, cigarillos, pipe tobacco, bidis or smokeless tobacco products. The present sequence represents a specifically claimed tobacco cytochrome P450 polynucleotide useful in marker assisted breeding methods. Note: The sequence data for this patent is also available in electronic format directly from the US patent office at <http://www.uspto.gov/sequence.html?DocID=20060041949>.

Sequence 1538 BP: 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other; XX

Query Match	23.7%	Score 410;	DB 15;	Length 1538;
Best Local Similarity	58.1%	Pred. NO. 1.5e-93;		
Matches 816;	Conservative	0;	Mismatches 555;	Indels 33

[illegible]

QY	869	ATCAAAGCATCCGGTTTTGTCGCGGAGCGAAACAACATTCGAAAGCTACA	928
DB	903	ATCAAAGCTATAATCTTTGACATGTTTGCCTGGAAACAGAGACTTCATCGTCAACAAAT	962
QY	929	GAATGGGTAAATGCGAGAGCTTGATGMAAAATCCAACTGAACTAAGAAAGACACAAGAAAGAA	988
DB	963	GTGTGGGCTATGGTGGAAATGGTGAAATAATCCAACTGTATTTGCGAAAGCTCAAGCAGAA	1022
QY	989	GTTAGACAAGTATTTGGTGGAAATGGGAAAAAGTTGATGAATCAAGATTTTCATGATTTGAAA	1048
DB	1023	GTAAGAGATGCAATTTAGAGAAAAAGAAACTTTTGTGATAAAATGATGTGGAGGAGCTAAAC	1082
QY	1049	TTCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATTACATCCTCGGTT---GTCCTTGATT	1105
DB	1083	TATCTAAAGTTAGTGCATTTAAAGAAACTCTAAGACTTCATCCACGGTTCCACTTTTGGCTC	1142
QY	1106	CCGAGGGAGTGTAGAGAAAAACAACGAATTTGATGGATATGAAATTCATCCGAACACTCGA	1165
DB	1143	CCAAGAGATGTAGGGAAGAGACAATATAAACGGCTACACTATTCCTGTAAAGACCAA	1202
QY	1166	ATTGTTCTGAATGCTTCGGCGATAGGAAGAGATCCTTAATACTTGTGCGGAACCTCGAAAG	1225
DB	1203	GTCATGTTTAATGTTTGGCGATTTGGGAAGAGATCCAAAATATTTGGGATGATGAGAAACT	1262
QY	1226	TTTAAACCCAGAAAGTTTAAAGATTGTCGAATTGATTATAAAGGACGACATTTTGAACTG	1285
DB	1263	TTTAAAGCCAGAGATTTTGACGAGTGCTCTAAGGATTTTGTGTGAATAATTTTGAATAT	1322
QY	1286	GTACCATTTGGTGAGGAAAAAGAAATGATGCTCGGCAATTACTTCAGCTATTACCAATTTG	1345
DB	1323	CTTCCATTTGGTGTGGAAGAGGATTTGTCAGGGATTTTCGTTGGTTTAGCTAATGCT	1382
QY	1346	GAGTATGTCATTATAAATCTATTATATCATTTTAAATTGGGAACTGGCCGATGGAATTTACA	1405
DB	1383	TATTTGCCATTGGCTCAATTACTTTATCACTTTGATTGGGAACTCCCCACTGGGAATCAA	1442
QY	1406	CCTCAAAACACTTGATATGACTGAAGCTATTGGGGGTGCTCTCAGGAAAAAAAATAGATCTT	1465
DB	1443	CCAAGCACTTGGACTTGACTGATGGTGGTGGAGTAATCTCCCGCTAGAAAAGTGACCTT	1502
QY	1466	AAGTTGATTCCTATTCCATATCAA	1489
DB	1503	TACTTGGTTCGACTCCTTATCAA	1526

RESULT 7
ADN10646

ID ADN10646 standard; cDNA; 1673 BP.

AC ADN10646;

XX
DT 15-JUL-2004 (first entry)

XX DE Nicotiana p450 enzyme encoding cDNA SEQ ID NO:231.

xx plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
kw
kw phenotype; gene; ss.

XX QS Nicotiana sp.

XX
PN
WO2004035745-A2.

XX
PD
29-APR-2004.

XX
PF 16-OCT-2003: 2003WO-US032722.XX
PR 16-OCT-2002: 2002US-0418933P.

PR 08-JUL-2003; 2003US-0485368P.
PR 18-SEP-2003; 2003US-0503989P.

XX
PA (USSM-) US SMOKELESS TOBACCO CO.

XX 16-OCT-2003; 2003WO-US032722.
PF
XX
PR 16-OCT-2002; 2002US-0418933P.
PR 08-JUL-2003; 2003US-0485368P.
PR 18-SEP-2003; 2003US-0503989P.
XX
PA (USSM-) US SMOKELESS TOBACCO CO.
XX
XX Xu D;
XX WPI; 2004-348441/32.
DR P-PSDB; ADN10625.
XX
PT New isolated nucleic acid molecules and encoded cytochrome P450 enzymes
from Nicotiana glauca, useful for altering plant phenotypes.
XX
PS Claim 1; SEQ ID NO 209; 198pp; English.
XX
CC The present sequence represents a plant p450 enzyme nucleic acid molecule
isolated from Nicotiana glauca. Also described: (i) an isolated p450 protein
from Nicotiana glauca; (ii) a transgenic plant comprising the nucleic acid
molecule described above; (iii) a method of producing the transgenic plant,
comprising: (i) operably linking the above nucleic acid molecule with a
promoter functional in the plant to create a plant transformational
vector; (ii) transforming the plant with the vector; (iii) selecting a
plant cell transformed with the transformation vector; and (iv)
regenerating a transformed plant from the transformed plant cell; and
(4) a method of selecting a plant containing the above nucleic acid
molecule, where the plant is analysed for the presence of the above
nucleic acid sequences. The p450 sequences have plant growth regulant
activity, and can be used in gene therapy. Compositions and methods from
the present invention are useful for altering plant phenotypes.
XX
SQ Sequence 1673 BP; 493 A; 309 C; 357 G; 514 T; 0 U; 0 Other;
Query Match 23.1%; Score 400.4; DB 12; Length 1673;
Best Local Similarity 57.7%; Pred. No. 4.4e-91;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
QY 113 CCTCAGGCGCATGGAAGTTTCTATCATAGTATCTTCCCTCATTTATTACTCACTTCT 172
DB 135 CCACAGGTCCTAGGAACTACCAATACCTAGGAGTATGCTTCATATGTT-----TGGT 188
QY 173 GATCTAGGCGCATGGAAGTTTCTAGAGCTTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
DB 189 GGAATACCAACCATGCTCTTAGAGATTTAGCCAAATAATAGGACCATTTATGACCTT 248
QY 233 CAAATTTGGCCAAATTTTCACTGTTGTCATTTCTTCACTGGAAGCAGCCAAAGAGTTATG 292
DB 249 CAATTAGGTGAAGTTTCTGCGGTTGGTTACTTCTCTGATACGCGCAAAAGAGATATTA 308
QY 293 AAAATCTCAGGCTGATGCTTCCGCCCAACGCTTATGCTTGGACCTGTTATGAGTCTT 352
DB 309 AAAATCTCAGCATGCTTTCGCTCTAGGCTTATGCTTTCCTGATACGCGCAAAAGAGATATTA 368
QY 353 TATAATCGGAAAGATGCTTGTGTTTGGCTTCATATGGAGATCATCTGGAGGCGAGATGAAGAAA 412
DB 369 TACAATAGTCTGATCTAGCTTTCGCCCTATGGCGACTATGGAGACAAATGCGTAAA 428
QY 413 ATTTGGATACCTGAATTTCTGAGTGGCCAAAGTTTCAATCTCCAGGTTTAATCCGAGAG 472
DB 429 ATATGTGCTTGGAGTGTCTAGTGGCAAGATGTTTCGGACATTTAGCTCTATTAGGCGG 488
QY 473 GAAGAAATCGGAGATGCCATCATCTTCCCTCGGTGGAAAGCGGATCTCCGGTCAATATT 532
DB 489 AATGAAGTCTTCGCTCTCATTTAATTTATCCGGTCAATCTTCTGGTGAACCTAATATGTT 548
QY 533 ACAAGATCATTTATGGCAATTATAATTTTCGATCATGATGAAGAACATCCGTTGGTAA---T 589
DB 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
QY 590 TGTAAGCAAAAGAAAGATGCTGAGTGTGGGATGTCAGTCAATGAGGCGCAGGAGT 649

DB 609 TTCAAGAGCAAGACAATAATTTATACAATAATTAAGAAGTGATATCTTTAGCAGAGGG 668
QY 650 TTTGCGACCGCAGACGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGGAGCTGAG 709
DB 669 TTTGATGTGGCTGACATATTTCCCTTCACTGAAGTTTCTTCACTGCTCAGTGGAAATGAAG 728
QY 710 TCAAAACCCAGGCGTTTTCATCAGGAGATTGACGATATACCTTGAAGAGATCTTTAATGAA 769
DB 729 GGTAAGATTATGAATGCACACCATTAAGGTAGATGTCATTGTTGAGAAATGTTCATCAATGAG 788
QY 770 CAC-----AAAGCCAAATAGCCTTTTGAAGCGGTAACCTTAATG 808
DB 789 CACAAGAAAAATCTTGCAATTTGGGAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACCTTCCAGTGCACGAGTCAAAAAACGAAAGC 868
DB 849 GATGTTCTTCAAGACTTATGAATGATGGAGCGCTTCAATTTCTCTATCACCACGACAAC 908
QY 869 ATCAAGCATCCGTTTTCGCAATGTTTACTCCGGGAGCGAAACAACTTTCGAAAGCTACA 928
DB 909 ATCAAGCATATAATTTTGTGATGTTTGTCTGCCGGGACAGAGACTTTCATCGTCAACAAT 968
QY 929 GAATGGGTAAATGCGCAGAGCTGATGAAAAATCCAACTGAACTAAGAAAAAGCACAAGAA 988
DB 969 GTGTGGGCTATGGTAGAAATGTTGAAAAATCCAGCCGTATTCGCGGAAAGCTCAACAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTGTGATGAATCAAGATTTTCATGATTTGAAA 1048
DB 1029 GTAGAGAGCATTTAGAGGAAAGAACTTTCGATGAAATGATGTGGAGGAGCTAAAC 1088
QY 1049 TCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTTACATCTCCGGTT---GTCTTGATT 1105
DB 1089 TACCTAAAGTTAGTAATAAAAGAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGTCTAGAGAAACAACGATTTGATGATGATGAAATTCATCCGAACTCGA 1165
DB 1149 CCAAGAGAAATGTAGGAAAGAGACAAATATAAACGCTACACTATTTCTGTAAAGACCAA 1208
QY 1166 ATTGTGTGAAATGCTTGGCGGATAGGAAAGAGATCCTAATATCTTGTGCGAACCTGGAAG 1225
DB 1209 GTCATGGTTAATGTTTGGGCTTTGGGAGAGATCCAAATATTTGGAATGACGCAAACT 1268
QY 1226 TTTAACCCAGAAAGGTTTAAAGATTGTCAAATTTAATTAAGGAGCAGACATTTGAACTG 1285
DB 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTTAAGGATTTTGTGTAATTAATTTTGAATAT 1328
QY 1286 GTACCATTTGTCGAGGAAAGAAAGATATGCTCTGCAATTTACTTTCAGCTATTACCAATTTG 1345
DB 1329 CTTCATTTGGTGGGAGGAGGATTTGCTCTGGGATTTGTTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGGAACTGGCGATGGAATACA 1405
DB 1389 TATTTGCCATTTGGCTCAATTAATCACTTTCGATTTGGAACTCCCTGCTGGAATCGAA 1448
QY 1406 CCTCAACACATTTGATATGATGAGCTATTTGGCGGTGCTCTCAGGAAAAAAATAGATCTT 1465
DB 1449 CCAAGCGCATTTGGACTTGAATGAGTTGGTGGAGTAACTGCGCGTAGAAAAAGTACGCTT 1508
QY 1466 AAGTTGATTTCTATTTCATATCAA 1489
DB 1509 TACTTGGTGGACTCTCTTATCAA 1532

RESULT 9

ADN10626
ID ADN10626 standard; cDNA; 1673 BP.
XX
AC ADN10626;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nicotiana p450 enzyme encoding cDNA SEQ ID NO:211.

XX plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
KW phenotype; gene; ss.
XX Nicotiana sp.
XX WO2004035745-A2.
XX 29-APR-2004.
XX 16-OCT-2003; 2003WO-US032722.
XX 16-OCT-2002; 2002US-0418933P.
XX 08-JUL-2003; 2003US-0485368P.
XX 18-SEP-2003; 2003US-0503989P.
XX (USSM-) US SMOKELESS TOBACCO CO.
XX Xu D;
XX WPI; 2004-348441/32.
XX P-PSDB; ADN10627.
XX New isolated nucleic acid molecules and encoded cytochrome P450 enzymes
XX from Nicotiana plants, useful for altering plant phenotypes.
XX Claim 1; SEQ ID NO 211; 198pp; English.
XX The present sequence represents a plant p450 enzyme nucleic acid molecule
XX isolated from Nicotiana. Also described: (1) an isolated p450 protein
XX from Nicotiana; (2) a transgenic plant comprising the nucleic acid
XX molecule described above; (3) a method of producing the transgenic plant,
XX comprising: (i) operably linking the above nucleic acid molecule with a
XX promoter functional in the plant to create a plant transformational
XX vector; (ii) transforming the plant with the vector; (iii) selecting a
XX plant cell transformed with the transformation vector; and (iv)
XX regenerating a transformation plant containing the above nucleic acid
XX molecule, where the plant is analysed for the presence of the above
XX nucleic acid sequences. The p450 sequences have plant growth regulant
XX activity, and can be used in gene therapy. Compositions and methods from
XX the present invention are useful for altering plant phenotypes.
XX Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;
Query Match 23.1%; Score 400.4; DB 12; Length 1673;
Best Local Similarity 57.7%; Pred. No. 4.4e-91;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
QY 113 CCTCAGGCGCCATGGAAGTTTCCTATCATAGGTAATCTTCCTCATTTTACTCACTTCT 172
DB 135 CCACAGGTCCATGGAACTACCAATACCTAGGAAGTATGCTTCATATGCT-----TGGT 188
QY 173 GATCTAGGCCATGAACGTTTATAGAGCTTGGCTCAAAATTTATGGACCTGTTATAGAGTCT 232
DB 189 GGACTACCAACACCATGCTCCTTAGAGATTTAGCCAAAATAATAGGACCACTTATGCACCTT 248
QY 233 CAAATTTGGCCAGTTTCAGCTGTTGCTATTTCTCAGCTGAGAGCCAGCCAAAGAGTTATG 292
DB 249 CAAATTAGGTGAAGTTTTCGCGTGTGGTGTACTTCTCTGTATACGGCAAAAGAATGATTA 308
QY 293 AAAAATCTCAGGCTGATGCCCTTCGCCCAACGCCCTATGCTCTTGGAGCGACAGATTTGTTT 352
DB 309 AAAAATCATGACATCGCTTTTGGGTCTAGGCTTAGCTTTTGGCCCGGAGATTTGCTGT 368
QY 353 TATAATCGGAAAGATGTCTTTGTTTGTCTCATATGAGATCATCTGGAGCGAGATGAAGAAA 412
DB 369 TACAATAGGTCTGATCTAGCTCTTTGGCCCTATGCGCACTATTGGAGACAAATGCGTAAA 428
QY 413 ATTTGGATCTTGAATTTCTGAGTCCAAAAGTTCAATCTCAGGTTTATCCGAG 472
DB 429 ATATGTGCTTGGAAAGTCTAGTGCCAAAGATGTTCCGACATTTAGCTCTATTAGCGG 488

QY 473 GAAGAAATGAGGATGCCATCACATTCCTCCGTTGAAAGCCGATCTCCGGTCAATATT 532
DB 489 AATGAAGTTCTTCGCTCATTAATTTATCCGGTCACTCTTGGTGAACTATTAATGTT 548
QY 533 ACAAGATCATTTATGGCATTATAATTTTCGATCATGATGAAGAACATCCGTTGGTAA--T 589
DB 549 ACGGAAAGGATCTTTTGTTCACAAGTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
QY 590 TGTNAGCAAAAGAAAGATTGCTGAGTGTGCGGATGCGATGCAATGAGGCGAGCAGAT 649
DB 609 TTCAAAGAGCAAGACAAATTTATACAACAAATTAAGAAGTGATATCTTTAGCAGGAGG 668
QY 650 TTTGGCACCGCAGAGCGCTTTCCGACGTGGAATTTACTTCACTATATCAATTTGAGCTGAG 709
DB 669 TTTGATGGCTGACATATTCCTTCACTGAGTTTCTTCATGCTGCTCAGTGGAAAGAAG 728
QY 710 TCAAAACCCAGCGCTTTGATCAGGAGATTGACGATATATCTTTGAAGAGATTTCTTAATGA 769
DB 729 GGTAAAGATTATGATGCACACCATAAAGGTAGATGCCATTTGTTGAGAATGTGTCATCAATGAG 788
QY 770 CAC-----AAAGCCAATNAGCCTTTTGAAGCGGATACCTTAATG 808
DB 789 CACAAGAAAAATCTTTGCAATTTGGGAAACTTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGAAAAAGTTCCAGTGCAGTGACAAACGAAAGC 868
DB 849 GATGTTCTTAAAGACTTATGATGATGGAGGCTTCAATTTCTTATACCAACGACAC 908
QY 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
DB 909 ATCAAAGCCATAATTTTTCGACATGTTTGTGCGGAGACAGACTTTCATCGTCAACAAAT 968
QY 929 GAATCGGTAAATGGCAGAGCTGATGAATTAATCCAACTGAACCTAAGAAAACGACAAAGAA 988
DB 969 GTGCGGCTATGTTAGAAATGTTGAAAAATCCAGCCGTTATCCGGAAGACTCAAGAGAA 1028
QY 989 GTTACACAAAGTATTTTGGTGAATGGAATGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
DB 1029 GTAAGAGAGCATTTTAGAGGAAAAAGAACTTTTCGATGAATAATGATGTGGAGGAGCTAAC 1088
QY 1049 TTTCTCAAGTTAGTGGTTTAAAGAACTCTAAGATTTACATCTCCCGTT---GTCTTGAT 1105
DB 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTTCACCCGCTTCCACTTTTGTCTC 1148
QY 1106 CCGAGGAGTGTAGAGAAACACACGAATTTGATGATGTAATTAATTCATCCGACACTCCA 1165
DB 1149 CCAAGAGAAATGTAGGGAAGAGACAAATATAAACCGCTACACTATTCTCTGTAAGACCAA 1208
QY 1166 ATTTGTTGAAATGCTTTGGCGGATAGGAAGAGATCCCTAATCTTGTGCGGAACCTGGAAG 1225
DB 1209 GTCATGTTAATGTTTGGCTTTGGGAGAGATCCAAATATTTGGAATGACGAGAACT 1268
QY 1226 TTTAAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGAGCAGCATTTGAAC 1285
DB 1269 TTTATGCCAGAGATTTTGGCAGTGTCTTAAGGATTTTGTGGTAAATAATTTTGAATAT 1328
QY 1286 GTACCATTGTTGCGAGGAAAGATATGCTCTGCGCATTTACTTACGATTTATCACTTTG 1345
DB 1329 CTTCCATTGTTGGCGGAAGGAGGATTTGCTCTGGGATTTCTGTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGGAACCTGGCCGATGAATTACA 1405
DB 1389 TATTTGCCATTGGCTCAATTACTATATCACTTCGATTTGGAACCTCCCTGCTGGAATCGAA 1448
QY 1406 CCTCAAAACATTTGATATGACTGAAGCTATTTGGCGTGTCTTCAGGAAAAAATAATGATCTT 1465
DB 1449 CCAAGCGACTTGGACTTCACTGAGTTGGTGGAGTAACTGCCCGTAGAAAAAGTACCTT 1508
QY 1466 AAGTTGATTCCTATTCCCATATCAA 1489
DB 1509 TACTTGGTTGCGACTCCTTTATCAA 1532

RESULT 10	Db	369	TACAATAGGCTGATCTAGCCCTTTTGCCCTATGCGGACTATTGGAGCAAAATGCGTAA	428
ADZ65334	Qy	413	ATTTGGATACCTTGAATTTCTGAGTGCCAAAGAAAGTTCAATCCTCCAGGTTAATCCGAGAG	472
ID ADZ65334 standard; cDNA; 1673 BP.	Db	429	ATATGTGCTTGGGAAGTGCTCAGTGCCAAAGAAATGTTGGACATTTAGCTCTATTAGGCGG	488
AC ADZ65334;	Qy	473	GAAGAAATGGAGGATGCCATCACATTCCTCGTTGCGAAAGCCGGATCTCCGGTCAATATT	532
14-JUL-2005 (first entry)	Db	489	AATGAAGTTCTCGTCTCATTAATTTTATCCGGTCATCTCTCTGGTGAACCTATTAAATGTT	548
Tobacco full length cDNA for cytochrome p450 clone D209-AH12.	Qy	533	ACAAAGATCATTTATGCAATTAATTTGCAATCATGATAAGAAATCCGTTGGTAA---T	589
Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;	Db	549	ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGATAGATCAGCGTTTGGGCAAGTG	608
nornicotine; transgenic plant; gene.	Qy	590	TGTAAGCAAAAGAAAGATTGCTGAGTGTTCGCCATGCGATCAATGAGGCGACGAGT	649
Nicotiana tabacum.	Db	609	TTCAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGATACTCTTAGCAGGAGGG	668
WO2005038033-A2.	Qy	650	TTTGGCCCGCAGAGCGCTTTCCGACGTGGAAATTTACTTCACTATATCATTTGGAGCTGAG	709
28-APR-2005.	Db	669	TTTGATGTGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGGAAATGAG	728
15-OCT-2004; 2004WO-US034065.	Qy	710	TCAAAACCCAGCGCTTTGTCATCAGAGATTGACGATATACCTTGAAGAGATTCTTAATGAA	769
16-OCT-2003; 2003US-00686947.	Db	729	GGTAAGATTGATGCACACCATAGGTAGTCCATTGTTGAGAATGTCATCAATGAG	788
29-APR-2004; 2004US-0568235P.	Qy	770	CAC-----AAAGCCAAATAGCCTTTTGAAGCGGATAAATCTTAATG	808
03-SEP-2004; 2004US-00934944.	Db	789	CACAAGAAATCTTTGCAATTTGGGAAACTAATGGAGCGTTAGGAGGTGAAGATTAAAT	848
(USM-) US SMOKELESS TOBACCO CO.	Qy	809	GATGTTCTATTGAATCTTCAAAAATATGGAACCGTTCCAGTGCCAGTCACAAACCAAGC	868
Xu D;	Db	849	GATGTTCTTCAAGCATTTATGAATGATGGAGCCCTTCAATTTCTTATCACCACACACAC	908
WPI; 2005-315717/32.	Qy	869	ATCAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA	928
P-PSDB; ADZ65335.	Db	909	ATCAAGCCATAATTTTGTGACATGTTGCTGCGGACAGAGACTTCATCGTCAACAATT	968
New nucleic acid molecule encoding cytochrome p450 enzymes in Nicotiana,	Qy	929	GAATGGGTAATGGCAGAGCTGATGAAATCCAACTGAATCAAGTAAAGAAAGCACAAAGAA	988
useful in developing tobacco plants with altered phenotypes.	Db	969	GTGTGGGCTATGTTAGAAATGTTGAAATCCAGCCGTAATTCGCGAAAGCTCAACAGCAA	1028
Disclosure; SEQ ID NO 231; 226pp; English.	Qy	989	GTTAGACAAGTATTTGGTGAATCGGAAAGTGTGATGAATCAAGATTTTCATGATTGAAA	1048
The invention relates to an isolated nucleic acid molecule (I) from	Db	1029	GTAAGAGAAGCATTTAGAGGAAAGAAACTTTTCATGAAATATGATGTGGAGGAGCTAAC	1088
Nicotiana, where the nucleic acid molecule comprising any of the 59	Qy	1049	TTCTTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCTTGATT	1105
nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to	Db	1089	TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC	1148
SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ	Qy	1106	CCGAGGGAGTGTAGAGAAACAAACGAAATTTGATGATATGAAATTCATCCGACACTCGA	1165
ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic	Db	1149	CCAAAGAAATGTAGGAAAGAGACAAATATAACCGCTACACTATTCTGTAAAGACCAAA	1208
acids of the invention encode cytochrome p450 enzymes whose expression is	Qy	1166	ATTGTTGTAATGCTTGGCGCATAGGAAGAGATCCTAATACTTGTGTCGAACTCGAAG	1225
induced by ethylene and/or plant senescence. Also included are a	Db	1209	GTCATGTTAATGTTTGGGCTTTGGGAAGAGATCCAAATATTGGAATGACCGCAAACT	1268
transgenic plant comprising (I), a method of producing a transgenic	Qy	1226	TTTAAACCAGAAAGGTTTTAAAGATTGTGCAATTTGATTAATAAGGAGCGACATTTGAAC	1285
plant, a method of selecting a plant containing a nucleic acid molecule	Db	1269	TTTATGCCAGAGAGATTTGACAGTGCTCTAAGGATTTTGTGGTAAATTAATTTGAATAT	1328
(where the plant is analyzed for the presence of nucleic acid sequence of	Qy	1286	GTACATTTGGTGAGGAAAGAAATATGTCCTGCGANTTACTTACAGCTATTACCAATTTG	1345
ADZ65402-	Db	1329	CTTCCATTTGGTGGGGAAGGAGGATTTGTCCTGGGATTTCTGTTGGCTTAGCTAATGCT	1388
Seq	Qy	1346	GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACCTGGCCGATGGAATTACA	1405
Query Match	Db	1389	TATTTGGCCATGGCTCAATTAATTAATCACTTCGATTTGGAAACTCCCTGCTGGAATCGAA	1448
Best Local Similarity	Qy	1406	CCTCAAAACATTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATGATCTT	1465
Matches 810; Conservative	Db	1449	CCAAGCGACTTGGACTTGACTTGAGTTGGTGGAGTAACTGCGCGCTAGAAAAAGTGACCTT	1508
113 CCTCCAGGCGCATGGAAGTTTCTCATCATAGTAATCTTCTCATTTTACTCTACTTCT	Qy	172		
135 CCACCAGGTCCTATGGAACCTACCAATACTAGGAAGTATGCTTCATATGTT-----TGGT	Db	188		
173 GATCTAGGCCATGAACGTTTATAGACCTTGGCTCAAAATTTATGGACCTTTATGATCTT	Qy	232		
189 GGACTACCAACCATGTCCTCTAGAGATTTAGCAAAAATATATGGACCACTTATGCACCTT	Db	248		
233 CAAATTGGCCAAAGTTTCTGTCATTTCTTCAGCTGGAAGCAGCAGCAAGAGGTTATG	Qy	292		
249 CAATTAGGTGAAGTTTCTCGGTTGTTGTTACTTCTCTGATACGCGCAAGAGTAATTA	Db	308		
293 AAAACTCAGGCTGATGCCCTTCGCCCAACGCCCTTATCGTCTTGGACGCAAGATTTGTTT	Qy	352		
309 AAAACTCATGACATCGCTTTTGGCTAGGCCCTTAGCCCTTTTGGCCCGGAGATTGTCGT	Db	368		
353 TATATCGAAGAGTCTTTGTTGCTTCATATGGAGATCACTGGGAGGAGATGAAGAA	Qy	412		

D_b 249 CAATTAGGTGAAGTTTCTGCGGTTGTGGTTACTTCTCTGATACGGCAAAGAGTATTA 308

```
Db 1389 TATTTCGCAATGGCTCAATTACTATATCACTTCGATTGGAAACTCCCTGCTGGAAATCGAA 1448
Qy 1406 CCTAAACACATTGATATCACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAAGTGGCTAGAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTCCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCCCTATCAA 1532

RESULT 12
ADZ65314
ID ADZ65314 standard; cDNA; 1673 BP.
XX AC ADZ65314;
XX DT 14-JUL-2005 (first entry)
XX DE Tobacco full length cDNA for cytochrome p450 clone D209-AA12.
XX KW Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
XX KW normicotine; transgenic plant; gene.
XX OS Nicotiana tabacum.
XX PN WO2005038033-A2.
XX XX
XX PD 28-APR-2005.
XX PF 15-OCT-2004; 2004WO-US034065.
XX PR 16-OCT-2003; 2003US-00686947.
XX PR 29-APR-2004; 2004US-0566235P.
XX PR 03-SEP-2004; 2004US-00934944.
XX PA (USSM-) US SMOKELESS TOBACCO CO.
XX PI Xu D;
XX WPI; 2005-315717/32.
XX DR P-PSDB; ADZ65315.
XX PT New nucleic acid molecule encoding cytochrome P450 enzymes in Nicotiana,
XX PS useful in developing tobacco plants with altered phenotypes.
XX PS Disclosure; SEQ ID NO 211; 226pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) from
XX CC Nicotiana, where the nucleic acid molecule comprising any of the 59
XX CC nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
XX CC SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
XX CC ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic
XX CC acids of the invention encode cytochrome p450 enzymes whose expression is
XX CC induced by ethylene and/or plant senescence. Also included are a
XX CC transgenic plant comprising (I), a method of producing a transgenic
XX CC plant, a method of selecting a plant containing a nucleic acid molecule
XX CC (where the plant is analyzed for the presence of nucleic acid sequence of
XX CC ADZ65402-
SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Query Match 23.1%; Score 400.4; DB 14; Length 1673;
Best Local Similarity 57.7%; Pred. No. 4.4e-91;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113 CCTCAGGCGCCATGGAAAGTTTCCTATCATAGTAATCTTCCTCATTTATCTCATCTTCT 172
Db 135 CCACCAGGTCCTATGGAAACTACCAATACTAGGAAGTATGCTTCATATGTT-----TGGT 188
Qy 173 GATCTAGGCCATGAACGTTTATAGAGCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
Db 173 GATCTAGGCCATGAACGTTTATAGAGCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
```

```
Db 189 GGACTACCAACCACATGCTCCTTAGAGATTTAGCCAAAAAATATGACCACCTTATGCACCTT 248
Qy 233 CAAATTGGCCAAAGTTTCAGCTGTTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTCGGTTGTGGTTACTTCTCTGATAGCGCAAAAGAGATATTA 308
Qy 293 AAAACTCAGGCTGATGCTTTCGCCCAACGCCCTATCGTCTTGGAGCGCACAGATCTGTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTTAGGCTTAGCCCTTTTGGCCCCGGAGATTTGCTGT 368
Qy 353 TATAATCGMAAGATGCTTCTTGTTCATATGGAGATCACTGGAGCGAGATGAAGAAA 412
Db 369 TACAATAGGTCTGATCTAGCCCTTTTGGCCCTATGGGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATACATTGAAATTTCTGAGTGCCCAAAAAGTTTCAATCTCCAGGTTTAAATCCGAGAG 472
Db 429 ATATGTGCTCTTGGAAAGTGTCTCAGTGCCAAAGATGTTCCGAGATTTTAGCTCTATTAGGCGG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTCGAAAGCGGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGTCCTCAITTAATTTTATCCGGTCATCTTCTGGTGAACCTATTAAATGTT 548
Qy 533 ACAAGATCAATTTATGTCATTTATATTTTCGATCATGATAAGAACAATCCGTTGGTAA---T 589
Db 549 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTCCTGAGTGTGCCGATGCGATCAATGAGCGCAGCGAGT 649
Db 609 TTCAAGAGCAAGACAAATTTATACAATAAATAAAGAGTGATACTCTTAGCAGGAGGG 668
Qy 650 TTTGGCCCGCAGACGCTTTTCCGACGTGGAAATTTACTTTCACTATATCATTTGGAGCTGAG 709
Db 669 TTTGATGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGAATGATGAG 728
Qy 710 TCAAAACCCAGCGCTTTGTCATCAGGAGATTGACGATATACTTGAAGAGATTCTTTAATGAA 769
Db 729 GGTAAAGATTGAATGCACACCAATAAGGTAGATGCCATTGTTGAGATGTCAATCAATGAG 788
Qy 770 CAC-----AAGCCCAATAGCCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CACAAGAAAAATCTTGCATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAATT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGAAACGTTCCAGTGCCAGTGACAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGCGCTTCAATTTCTCTATCACCAACCAACAC 908
Qy 869 ATCAAAGCATCCGTTTTTGCATAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAAGCCATAATTTTGGACATGTTTCTGCGGGACAGAGACTTCATCGTCAACAATT 968
Qy 929 GAATGGGTAATGGCAGAGCTGATGAAAAATCCAATGAACTGAAGAAAAAGCACAAGAGAA 988
Db 969 GTGTGGGCTATGTTAGAAATGTTGAAAAATCCAGCCGTTATTCGCGAAAGCTCAAGCAGAA 1028
Qy 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTCATGATTTGAAA 1048
Db 1029 GTAAGAGAAGCAATTTAGAGGAAAAAGAACTTTTCGATGAAAATGATGTGGAGAGCTAAAC 1088
Qy 1049 TCTTCAAAGTTAGTGGTTTAAAGAAACTCTAAGATTACATCTCCCGGTT---GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAAATAAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGGTGATAGAAAAACAAACGAATTTGATGATATGAATTAATTCATCCGAACTCGA 1165
Db 1149 CCAAGAGAATGTAGGGAAGAGACAAATATATAACCGCTACACTATTCTCTGTTAAAGACCAA 1208
Qy 1166 ATTTGTTGTAATGCTTTGGGCGATAGGAAGATCCTTAATACCTTGGTCGGAACCTCGAAAG 1225
Db 1209 GTCTGGTTAATGTTTGGGCTTTTGGGAAGAGATCCAAATATTTGAAATGACGAGAACT 1268
Qy 1226 TTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTAATAAAGGAGCGACATTGTA 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTAAGGATTTTGGTAAATAATTTTGAATAT 1328
```


Db 1029 GTAAGAGAGCAATTTAGAGGAAAGAAACTTTTCGATGAAATGATGTGGAGAGCTAAAC 1088
Qy TTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCTCCGGTT---GTCCTTGATT 1105
Db TTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCTCCGGTT---GTCCTTGATT 1148
Qy 1106 CCGAGGAGTGTAGAGAAACACAGCAATTTGATGATATGAAATTCATCCGACACTCGA 1165
Db 1149 CCAAGAGAATGTAGGAAAGAGACAAATATAAACGGCTACACTATTTCCTGTAAAGACCAA 1208
Qy 1166 ATTGTTGTGAATGCTTTGGCGGATAGAGAGATCCTTAATCTTGTTCGGAACCTGGAAG 1225
Db 1209 GTCATGGTTAATGTTTGGCGTTTGGGAAGAGATCCAAATATTTGGAATGACGCAAACT 1268
Qy 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGATTTGAGCAGTGCTCTAAGGATTTTGTGGTAATATTTGAAATAT 1328
Qy 1286 GTACATTTGGTGCAGGAAAGAAATATGTCCTGGCATTACTTCAGCTATTACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGCGGAGGAGGATTTGTCCTGGGATTTGCTTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATATCAATTTTAAATGGGAAGTGGCGATGGAAATACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTTACTATATCACTTCGATTGGAATCTCCCTGCTGGAAATCGAA 1448
Qy 1406 CCTCAACACATTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAGAAATAGATCTT 1465
Db 1449 CCAAGCGACTGGACTTGACTGAGTGGTTGGAGTAACCTGCCGCTAGAAAGTGAACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532

RESULT 14

ID ADZ65859 standard; cdna; 1673 BP.

AC ADZ65859;

DT 14-JUL-2005 (first entry)

DE Tobacco cytochrome P450 enzyme cdna #105.

KW Enzyme engineering; cytochrome P450; gene; ss.

OS Nicotiana tabacum.

PN WO2005038018-A2.

PD 28-APR-2005.

PP 15-OCT-2004; 2004WO-US034218.

PR 16-OCT-2003; 2003US-00686947.

PR 29-APR-2004; 2004US-0566235P.

PR 17-SEP-2004; 2004US-00943507.

XX (USSM-) US SMOKELESS TOBACCO CO.

XX Xu D;

XX WPI; 2005-315709/32.

XX P-PSDB; ADZ65860.

XX New isolated nucleic acid molecule from Nicotiana, useful for altering plant phenotypes, thus producing a transgenic plant having reduced levels of nicotine.

XX Disclosure; SEQ ID NO 209; 203pp; English.

XX The invention relates to an isolated nucleic acid molecule from

CC Nicotiana, encoding a protein. The invention also relates to a transgenic plant comprising the nucleic acid molecule, a method of producing a transgenic plant comprising operably linking the nucleic acid molecule with a promoter functional in the plant to create a plant transformational vector, transforming the plant with the plant transformational vector, selecting a plant cell transformed with the transformational vector and regenerating a transformation plant from the transformed plant cell, a method of selecting a plant containing a nucleic acid molecule, a method of increasing or decreasing nicotine levels in a plant by operably linking the nucleic acid molecule with a promoter functional in the plant, a tobacco product having reduced amounts of nicotine levels, the tobacco product comprising tobacco from the plant, a tobacco leaf having reduced amounts of nicotine levels and a method of isolating a gene from a plant using the isolated nucleic acid. In producing a transgenic plant, the plant has reduced levels of nicotine. The tobacco product is selected from cigarettes, cigars, pipe tobacco, snuff, chewing tobacco, products blended with the tobacco product and their mixtures. The nucleic acid molecule is useful for altering plant phenotypes, thus producing a transgenic plant having reduced levels of nicotine. This sequence represents cdna encoding a tobacco cytochrome P450 enzyme of the invention.

SQ Sequence 1673 BP; 493 A; 309 C; 357 G; 514 T; 0 U; 0 Other;

Query Match 23.1%; Score 400.4; DB 14; Length 1673;

Best Local Similarity 57.7%; Pred. No. 4.4e-91;

Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113 CCTCAGGGCCATGGAGTTTCTTATCATAGTAATCTTCTCATTTTATTACTCACTTCT 172

Db 135 CCACAGGTCCATGGAACTACCAATACTAGAGATGCTTCATATGTT-----TGCT 188

Qy 173 GATCTAGGCCATGAACGTTTGTAGAGCCTTGCTCAAATTTATGGACCTGTTATGAGTCT 232

Db 189 GGACTACCACACCATGCTCTTAGAGATTTAGCCAAAATATGGACCACTTATGCACCTT 248

Qy 233 CAATTTGCCCAAGTTTCAGCTGTTGTCATTTTCTCAGCTGAAGCAGCAAGAGGTTATG 292

Db 249 CAATTAGGTGAAGTTTCTGCGGTTGTGGTTTACTTCTCTGATACGGCAAGAGATATTA 308

Qy 293 AAAACTCAGGCTGATGCCTTCGCCCAACGCCCTATGCTTTGGAGCAGACAGTTGTGTT 352

Db 309 AAACTCATGACATCGCTTTTGGCTCTAGGCTAGCCCTTTTGGCCCGGAGATTCTCTGT 368

Qy 353 TATAATCGGAAAGATGCTTGTGTTTCATATGAGATCACTGGAGCAGATGAAGAA 412

Db 369 TACAATAGGCTGATCTAGCCCTTTTGGCCCTATGGCGACTATTGGAGACNAATGCGTAAA 428

Qy 413 ATTTGGATACCTTGAATTTCTGAGTGCCAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472

Db 429 ATATGTGTTTGGAAAGTCTCAGTGCCAAAGATGTTCCGACATTTAGCTCTATTAGGCGG 488

Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCCGTTTGGAAAGCCGGATCTCCGTCATATT 532

Db 489 AATGAAGTTCTTCGTTCTCATTAATTTTATCCGTCATCTTCTGTGGAACCTATTATGTT 548

Qy 533 ACAAGATCATTTATGGCATTATTAATTTGATCATGATGAAGAACATCCGTTGGTAA---T 589

Db 549 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608

Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTGCGGATGCGATCAATGAGGACGACGAGT 649

Db 609 TTCAAAGAGCAAGCAAAATTTATCAACTAATTAAGAAGTGATCTCTTAGCAGAGGG 668

Qy 650 TTTGGCACCCGACGAGCTTTTCCGACGTTGGAATTAATCTTCACTATATCATTTGAGCTGAG 709

Db 669 TTTGATGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGTCAGTGGAAATGAAG 728

Qy 710 TCAAAACCCAGCGTTTGCATCAGAGATTGACATATACTTTGAAGAGATTTCTTATGAA 769

Db 729 GGTAAGATTGAATGCACACCATTAAGGTAGATGCCATTTGTTGGAATGTCTCATCAATGAG 788

Qy 770 CAC-----AAAGCCAATAAGCCTTTTGAAGCGGATAACTTAATG 808

Db 789 CACAAGAAAAATCTTGCATTTGGAAAACTAATGGAGCGTTAGGAGTGAAGATTTAATT 848
QY 809 GATGTTCTTAAATCTTCAAAAAATGGAAAGTTCCAGTGCAGTGACAAACGAAGC 868
Db 849 GATGTTCTTAAATCTTCAAGACTATGATGATGGAGGCTTCAATTTTCCTATCAACGACAAC 908
QY 869 ATCAAGACATCCGTTTTCGAAATGTTTACTGCGCGGAGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAGACTATAATTTTGCATGTTTGTGCGCGGACAGAGACTTCATCGTCACAAATT 968
QY 929 GAATGGGTAAATGCGAGCTGATGAAAAATCCAACTGAACCTAAGAAAGCACAAGAGAA 988
Db 969 GTGTGGCTATGTAAGAAATGGTGAATAATCCAGCCGTATTTCGCGAAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTTCGTGAATGGGAAAGTTGATGAATCAAGATTTCATGATTTGAAA 1048
Db 1029 GTAAGAGAAGTATTTAGAGGAAAGAACTTTTCGATGAATAATGATGAGGAGCTAAAC 1088
QY 1049 TTCTTCAAGTTAGTGTGTTAAAGAACTCTAAGATTACATCCTCCGGTT--GCTTTGATT 1105
Db 1089 TACCTAAGTTAGTATAAAGAAACTCTAAGACTTCATCCACCGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGTTAGAGAAACAAACACGAATTTGATGGATATCAAAATTCATCCGAACCTCGA 1165
Db 1149 CCAAGAGAATGTAGGAAGAGACAAATATAAACGGCTACACTATTCTCTGTAAGACCAA 1208
QY 1166 ATGTTGTGAATGCTTGGCGGATAGGAAGATCTTAATCTGTCGGAACCTGGAAG 1225
Db 1209 GTCATGGTTAATGTTTGGGCTTTGGGAAGAGATCCAAATAATTTGGAATGACGAGAACT 1268
QY 1226 TTTAACCCAGAAAGTTTAAAGATTGTGCAATTGATTATAAGGGACGACATTTGAAC TG 1285
Db 1269 TTTATGCCAGAGAGATTTGACAGTCTCTAAGGATTTTGTGGTAAATTTGATAT 1328
QY 1286 GTACATTTTGGTCAGGAAAAAGAAATGATGCTCGGCATTTACTTTCAGCTATTACCAATTTG 1345
Db 1329 CTTCATTTTGGTGGCGAAGGAGGATTTGCTCGGATTTCTGCTGCTAGCTAATGCT 1388
QY 1346 GAGTATGCTATTATAATCTATTATATCATTTTAAATGGGAACTGCGCGATGAATTACA 1405
Db 1389 TATTTGCCAATGGCTCAATTACTATATCACTTCGATTGGAATCCCTGCTGGAATCGAA 1448
QY 1406 CTTCAAAACATTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCGACTTGGACTTGAATGATGTTGGTGGATGAATCTGCGCTAGAAAAAGTGACCTT 1508
QY 1466 AAGTTGATTCCTATTCATATCAAA 1489
Db 1509 TACTTGGTTGCGACTCCTTATCAA 1532

RESULT 15

ADZ65881
ID ADZ65881 standard; cDNA; 1673 BP.
XX
AC ADZ65881;

DT 14-JUL-2005 (first entry)

DE Tobacco cytochrome P450 enzyme cDNA #116.

KW Enzyme engineering; cytochrome P450; gene; ss.

OS Nicotiana tabacum.

XX WO2005038018-A2.

XX 28-APR-2005.

PF 15-OCT-2004; 2004WO-US034218.

PR 16-OCT-2003; 2003US-00686947.

PR 29-APR-2004; 2004US-0566235P.
PR 17-SEP-2004; 2004US-00943507.
XX (USM-) US SMOKELESS TOBACCO CO.
XX Xu D;
XX
DR WPI: 2005-315709/32.
DR P-PSDB; ADZ65882.
XX
PT New isolated nucleic acid molecule from Nicotiana, useful for altering
PT plant phenotypes, thus producing a transgenic plant having reduced levels
XX of nicotine.
PS Disclosure; SEQ ID NO 231; 203pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule from
CC Nicotiana, encoding a protein. The invention also relates to a transgenic
CC plant comprising the nucleic acid molecule, a method of producing a
CC transgenic plant comprising operably linking the nucleic acid molecule
CC with a promoter functional in the plant to create a plant
CC transformational vector, transforming the plant with the plant
CC transformational vector, selecting a plant cell transformed with the
CC transformed plant cell, a method of selecting a plant containing a
CC nucleic acid molecule, a method of increasing or decreasing nicotine
CC levels in a plant by operably linking the nucleic acid molecule with a
CC promoter functional in the plant, a tobacco product having reduced
CC amounts of nicotine levels, the tobacco product comprising tobacco
CC from the plant, a tobacco leaf having reduced amounts of nicotine
CC levels and a method of isolating a gene from a plant using the isolated
CC nucleic acid. In producing a transgenic plant, the plant has reduced
CC levels of nicotine. The tobacco product is selected from cigarettes,
CC cigars, pipe tobacco, snuff, chewing tobacco, products blended with the
CC tobacco product and their mixtures. The nucleic acid molecule is useful
CC for altering plant phenotypes, thus producing a transgenic plant having
CC reduced levels of nicotine. This sequence represents cDNA encoding a
CC tobacco cytochrome P450 enzyme of the invention.
XX
SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Query Match 23.1%; Score 400.4; DB 14; Length 1673;
Best Local Similarity 57.7%; Pred. No. 4.4e-91;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTCATCATAGTAAATCTTCCTCATTTATTACTCACTTCT 172
Db 135 CCACCAGGTCATGGAACACTACCAATACTAGGAAGTATGCTTCATATGCT-----TGCT 188
QY 173 GATCTAGGCCATGAACGTTTATAGAGCTTGGCTCAAATTTATGGACCTGTTATGAGTCTT 232
Db 189 GGACTACACACCATTGCTCTTAGAGATTAGCCAAATAATATGGACCACTTATGACCTT 248
QY 233 CAAATTGGCCAGTTTTCAGCTGTTGCTCATTTCTCAGCTGAAGACGCCAAGAGGTTATG 292
Db 249 CAAATTAGGTGAAGTTTCTGGGTTGTTGTTACTTCTCTGTATACGCCAAGAAGATTATTA 308
QY 293 AAAACTCAGGCTGATGCTTTCGCCCAACGCCCTTATCGTCTTGGACGCAAGATTGTTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTTCGCTTAGGCCCTAGCCCTTTTGGCCCCCGAGATTGCTCT 368
QY 353 TATAATCGAAAGATGCTGTTGTTCTCATATGAGATCACTGGAGGCGAGATGAAGAA 412
Db 369 TACAATAGTCTGATCTAGCCCTTTGGCCCTATGGGACTATTGGAGACAATATGCGTAAA 428
QY 413 ATTTGGATATCTTGAAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGTGCTTGGAAAGTCTCAGTGCACGAAGATGTTGCGACATTTTAGCTCTATTAGCGG 488
QY 473 GAAGAAATGGAGGATGCCATCACTTCTCCGTTGGAAGCCGATCTCCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCTCGTCTCATTAATTTTATCCGGTCTATCTTCTGTTGGAACCTATTAA 548

QY 533 ACAAGATCATTTATGGCATTATAATTTCGATCATGATAAGAACATCCGTTGGTAA---T 589
|||
Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
|||
QY 590 TGTAGCAAAAGAAAGATTGCTGAGTGTTCGCGATGCGATCAATGAGGCGACGAGT 649
|||
Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAAATTAAGAAGTGATACTCTTAGCAGGAGG 668
|||
QY 650 TTTGGCACCGCAGAGCGCTTTCCGACGTGGAAATTACTTCACTATATCATTTGAGCTGAG 709
|||
Db 669 TTTGATGTGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGAATGAAG 728
|||
QY 710 TCAAAACCCAGCGGTTTGATCAGGAGATTGACGATATACCTTTGAAGAGATTCTTAATGAA 769
|||
Db 729 GGTAAGATTATGATGACACCATTAAGGTAGATGCCATTGTTGAGAATGTCATCAATGAG 788
|||
QY 770 CAC-----AAAGCCATAAGCCTTTTGAAGCGGATAACTTTAATG 808
|||
Db 789 CAAAGAAAAATCTTGGCAATTGGGAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
|||
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTCCAGTGCCAGTGACAAACGAAAGC 868
|||
Db 849 GATGTTCTTAAGACTTATGAATGATGGAGGCTTCAATTTCTCTATCACCACGACAAAC 908
|||
QY 869 ATCAAAGCATCCGTTTGGCAAAATGTTTACTGCGGGAGCGGAAACAACTTCGAAAGCTACA 928
|||
Db 909 ATCAAAGCCATAATTTTGCATGTTTGTGCTGGGACAGAGACTTCATCGTCAACAATT 968
|||
QY 929 GAATGGTAAATGCGAGAGCTGATGAAATCCAACTGAACCTAAGAAAGCACAAGAA 988
|||
Db 969 GTGTGGCTATGTTAGAAATGGTGAATAATCCAGCCGTATTCGGAAAGCTCAAGCAGAA 1028
|||
QY 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
|||
Db 1029 GTAAGAGAGCATTTAGAGGAAAGAAACTTTTCGATGAAATGATGTGGAGGAGCTAAAC 1088
|||
QY 1049 TTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCCGTT---GTCTTGATT 1105
|||
Db 1089 TACCTAAAGTTAGTAAATAAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
|||
QY 1106 CCGAGGAGTGTAGAGAAACACACGATTTGATGGATGAAATTCATCCGAACTCGA 1165
|||
Db 1149 CCAAGAGAAATGTAGGAAGAGACAAATATAAACGGCTACACTATTCCTGTAAGACCAA 1208
|||
QY 1166 ATTGTTGTAATGCTTTGGCGATAGGAGAGATCCTAATACCTTGGTCGGAACCTGGAAG 1225
|||
Db 1209 GTCATGGTTAATGTTTGGGCTTTGGAGAGATCCAAATAATTGGAATGACGCAGAACT 1268
|||
QY 1226 TTTAACCAGAAAGGTTTAAAGATTGTGCAATTGATTATTAAGGGACGACATTTGAACTG 1285
|||
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTTGTGGTAATAATTTTGAATAT 1328
|||
QY 1286 GTACCATTTGGTGCAAGAAAAGAAATATGCTCTGGCATTACTTCAGCTATTACCAATTTG 1345
|||
Db 1329 CTTCCATTTGGTGGCGAAGGAGGATTTGTCTGGGATTTCTGTTGGCTTAGCTAATGCT 1388
|||
QY 1346 GAGTATGTCATTATAAATCTATTATATCAATTTAATTGGAACTGGCCGATGGAATTACA 1405
|||
Db 1389 TATTTGCCATTGGCTCAATTAATACTATATCACTTCGATTGGAAACTCCCTGCTGGAATCGAA 1448
|||
QY 1406 CTTCAAACTTTGATATGATGACTAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
|||
Db 1449 CCAAGCGACTTGGACTTGACTGAGTTGGTGGAGTAACGCGCTAGAAAAAAGTGACCTT 1508
|||
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
|||
Db 1509 TACTTGGTGGCACTTCTTATCAA 1532
|||

GenCore version 5.1.1.8
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 09:33:07 ; Search time 2310 seconds
 (without alignments)
 9218.386 Million cell updates/sec

Title: US-10-759-813-1
 Perfect score: 1733
 Sequence: 1 gataaagggaataggagc.....tcttttcaaatccgaaaaa 1733

Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 18992170 seqs, 6143817638 residues
 37784340

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733	100.0	1733	3	US-09-909-566C-1
2	1733	100.0	1733	8	US-10-759-813-1
3	410	23.7	1538	10	US-10-686-947-153
4	410	23.7	1538	10	US-10-943-507-153
5	410	23.7	1538	11	US-10-934-944-153
6	410	23.7	1538	16	US-11-116-881A-162
7	400.4	23.1	1673	8	US-10-686-947-209
8	400.4	23.1	1673	8	US-10-686-947-211
9	400.4	23.1	1673	8	US-10-686-947-231
10	400.4	23.1	1673	10	US-10-943-507-206
11	400.4	23.1	1673	10	US-10-943-507-208
12	400.4	23.1	1673	11	US-10-934-944-209
13	400.4	23.1	1673	11	US-10-934-944-211
14	400.4	23.1	1673	11	US-10-934-944-231
15	400.4	23.1	1673	16	US-11-116-881A-218
16	400.4	23.1	1673	16	US-11-116-881A-220
17	400.4	23.1	1673	16	US-11-116-881A-240

18	398.8	23.0	1673	8	US-10-686-947-213	Sequence 213, App
19	398.8	23.0	1673	10	US-10-943-507-210	Sequence 210, App
20	398.8	23.0	1673	11	US-10-934-944-213	Sequence 213, App
21	398.8	23.0	1673	16	US-11-116-881A-222	Sequence 222, App
22	393.2	22.7	1610	8	US-10-686-947-197	Sequence 197, App
23	393.2	22.7	1610	8	US-10-686-947-261	Sequence 261, App
24	393.2	22.7	1610	10	US-10-943-507-194	Sequence 194, App
25	393.2	22.7	1610	10	US-10-943-507-256	Sequence 256, App
26	393.2	22.7	1610	11	US-10-934-944-197	Sequence 197, App
27	393.2	22.7	1610	11	US-10-934-944-261	Sequence 261, App
28	393.2	22.7	1610	16	US-11-116-881A-206	Sequence 206, App
29	393.2	22.7	1610	16	US-11-116-881A-270	Sequence 270, App
30	388.8	22.4	1566	8	US-10-686-947-191	Sequence 191, App
31	388.8	22.4	1566	10	US-10-943-507-188	Sequence 188, App
32	388.8	22.4	1566	11	US-10-934-944-191	Sequence 191, App
33	388.8	22.4	1566	16	US-11-116-881A-200	Sequence 200, App
34	388.8	22.4	1576	8	US-10-686-947-149	Sequence 149, App
35	388.8	22.4	1576	10	US-10-943-507-149	Sequence 149, App
36	388.8	22.4	1576	11	US-10-934-944-149	Sequence 149, App
37	388.8	22.4	1576	16	US-11-116-881A-158	Sequence 158, App
38	388.8	22.4	1581	8	US-10-686-947-151	Sequence 151, App
39	388.8	22.4	1581	10	US-10-943-507-151	Sequence 151, App
40	388.8	22.4	1581	11	US-10-934-944-151	Sequence 151, App
41	388.8	22.4	1581	16	US-11-116-881A-160	Sequence 160, App
42	388.8	22.4	1664	8	US-10-686-947-199	Sequence 199, App
43	388.8	22.4	1664	10	US-10-943-507-196	Sequence 196, App
44	388.8	22.4	1664	11	US-10-934-944-199	Sequence 199, App
45	388.8	22.4	1664	16	US-11-116-881A-208	Sequence 208, App

ALIGNMENTS

RESULT 1

US-09-909-566C-1

; Sequence 1, Application US/09909566C

; Publication No. US20030066103A1

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Edgar B

; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-

; FILE REFERENCE: BBI465 US NA

; CURRENT APPLICATION NUMBER: US/09/909,566C

; PRIOR FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: 60/219833

; PRIOR FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1733

; TYPE: DNA

; ORGANISM: Euphorbia lagascae

US-09-909-566C-1

Query Match	100.0%	Score 1733;	DB 3;	Length 1733;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1733;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCATAAAGGAAATGGAGCAGAAAAATCTCTCTTTCCGAGCATTTTAAATAGTTTCT	60	
Db	1	GCATAAAGGAAATGGAGCAGAAAAATCTCTCTTTCCGAGCATTTTAAATAGTTTCT	60	
QY	61	GCTTCCTTTATCTTAGTAGTATCATGAGGTTGGAGAGAACAGAAATCCACCTCCAG	120	
Db	61	GCTTCCTTTATCTTAGTAGTATCATGAGGTTGGAGAGAACAGAAATCCACCTCCAG	120	
QY	121	GCATGGAAGTTTCTTATCATAGGTAATCTTCTCTATTTATTTACTTCTGATCTAGG	180	
Db	121	GCATGGAAGTTTCTTATCATAGGTAATCTTCTCTATTTATTTACTTCTGATCTAGG	180	
QY	181	CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGAGCTGTATGAGTCTTCAATGG	240	
Db	181	CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGAGCTGTATGAGTCTTCAATGG	240	

```
QY 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATGAAAACTCA 300
DB |||||
QY 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATGAAAACTCA 300
DB |||||
QY 301 GGCTGATGCTTTCGCCCCAACCGCCCTATCGTCTTGGACGACAGATGTGTTTTATATATCG 360
DB |||||
QY 301 GGCTGATGCTTTCGCCCCAACCGCCCTATCGTCTTGGACGACAGATGTGTTTTATATATCG 360
DB |||||
QY 361 GAAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGGCGAGATGAAGAAAAATTTGGAT 420
DB |||||
QY 361 GAAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGGCGAGATGAAGAAAAATTTGGAT 420
DB |||||
QY 421 ACTTGAATTTCTGAGTGCCAAAAAGTTCAATCCCTCCAGGTTAAATCCGAGAGGAGAAAT 480
DB |||||
QY 421 ACTTGAATTTCTGAGTGCCAAAAAGTTCAATCCCTCCAGGTTAAATCCGAGAGGAGAAAT 480
DB |||||
QY 481 GGAGATGCGCATCAATTCCTCCGTTGAAAAGCCGGATCTCCGGTCAATATTAACAAGAT 540
DB |||||
QY 481 GGAGATGCGCATCAATTCCTCCGTTGAAAAGCCGGATCTCCGGTCAATATTAACAAGAT 540
DB |||||
QY 541 CATTTATGCGCATTAATTTGCGATCATGATAGAACATCCGTTGGTAATTTGTAAGCAAAA 600
DB |||||
QY 541 CATTTATGCGCATTAATTTGCGATCATGATAGAACATCCGTTGGTAATTTGTAAGCAAAA 600
DB |||||
QY 601 AGAAAGATTGCTGAGTGTGCGGATGCGCATCAATGAGGCGAGCGAGTTTGGCACCGC 660
DB |||||
QY 601 AGAAAGATTGCTGAGTGTGCGGATGCGCATCAATGAGGCGAGCGAGTTTGGCACCGC 660
DB |||||
QY 661 AGACGCTTTTCGAGCGTGAATTAATCTTCACTATATCATTTGGAGCTGAGTCAAAACCCAG 720
DB |||||
QY 661 AGACGCTTTTCGAGCGTGAATTAATCTTCACTATATCATTTGGAGCTGAGTCAAAACCCAG 720
DB |||||
QY 721 GCGTTTGATCAGAGAGATTGACGATATCTTGAAGAGATCTTAAATGAACACAAAGCCAA 780
DB |||||
QY 721 GCGTTTGATCAGAGAGATTGACGATATCTTGAAGAGATCTTAAATGAACACAAAGCCAA 780
DB |||||
QY 781 TAAGCCTTTTGAAGCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAA 840
DB |||||
QY 781 TAAGCCTTTTGAAGCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAA 840
DB |||||
QY 841 CGTTCAGTCCAGTGCACAAAGCAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGC 900
DB |||||
QY 841 CGTTCAGTCCAGTGCACAAAGCAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGC 900
DB |||||
QY 901 CGGGAGCGAAACAACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCC 960
DB |||||
QY 901 CGGGAGCGAAACAACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCC 960
DB |||||
QY 961 AACTGAACCTAAGAAAAGCAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAAAT 1020
DB |||||
QY 961 AACTGAACCTAAGAAAAGCAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAAAT 1020
DB |||||
QY 1021 TGATGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAACTCTAAG 1080
DB |||||
QY 1021 TGATGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAACTCTAAG 1080
DB |||||
QY 1081 ATTACATCTCCGGTTGCTTCATTTCCGAGGAGTGTAGAGAAACAAACGAAATTTGATGG 1140
DB |||||
QY 1081 ATTACATCTCCGGTTGCTTCATTTCCGAGGAGTGTAGAGAAACAAACGAAATTTGATGG 1140
DB |||||
QY 1141 ATATGAAATTTATCCGAACACTCGAATTTGTTGTAATGCTTGGCGGATAGGAAGATCC 1200
DB |||||
QY 1141 ATATGAAATTTATCCGAACACTCGAATTTGTTGTAATGCTTGGCGGATAGGAAGATCC 1200
DB |||||
QY 1201 TAATACTTGGTCGGAACCTGGAAGTTTAAACCCAGAAAGTTTAAAGATTTGCAATGCA 1260
DB |||||
QY 1201 TAATACTTGGTCGGAACCTGGAAGTTTAAACCCAGAAAGTTTAAAGATTTGCAATGCA 1260
DB |||||
QY 1261 TTATAAAGGGAGCAATTTGAACCTGGTACCATTTGGTCAGGAAAAAGAAATATGTCCTGG 1320
DB |||||
QY 1261 TTATAAAGGGAGCAATTTGAACCTGGTACCATTTGGTCAGGAAAAAGAAATATGTCCTGG 1320
DB |||||
```

```
QY 1321 CATTACTTCAGCTATTACCAATTTTGGAGTATGTCAATTAATAAATCTATTATATCATTTTAA 1380
DB |||||
QY 1321 CATTACTTCAGCTATTACCAATTTTGGAGTATGTCAATTAATAAATCTATTATATCATTTTAA 1380
DB |||||
QY 1381 TTGGGAACCTGCCCCGATGGAATTTACACCTCAAACTCTGATGACTGAAGCTATTGGCGG 1440
DB |||||
QY 1381 TTGGGAACCTGCCCCGATGGAATTTACACCTCAAACTCTGATGACTGAAGCTATTGGCGG 1440
DB |||||
QY 1441 TGCTCTCAGGAAAAAATAGATCTTAAAGTTGATTTCCATATCCATATCAAGTTAGCTTAGG 1500
DB |||||
QY 1441 TGCTCTCAGGAAAAAATAGATCTTAAAGTTGATTTCCATATCCATATCAAGTTAGCTTAGG 1500
DB |||||
QY 1501 CTCAAAATATTCTTGTGATTAATAGAGGTTGAAATATATATATAATAAATCTTTAATTAACG 1560
DB |||||
QY 1501 CTCAAAATATTCTTGTGATTAATAGAGGTTGAAATATATATAATAAATCTTTAATTAACG 1560
DB |||||
QY 1561 ATGTTCTTAATATGTTTGGGTGAGTTAATAGGTTTTCCACCGCATCATATAAGTAGCCT 1620
DB |||||
QY 1561 ATGTTCTTAATATGTTTGGGTGAGTTAATAGGTTTTCCACCGCATCATATAAGTAGCCT 1620
DB |||||
QY 1621 TCTTTGATGATGGGTTAGATTAATAGTGTGTTGGTTGGATTTTATAGATGGGTTAAAT 1680
DB |||||
QY 1621 TCTTTGATGATGGGTTAGATTAATAGTGTGTTGGTTGGATTTTATAGATGGGTTAAAT 1680
DB |||||
QY 1681 GATTGGATGATTAATAATAAATTTGAAATGTTTCTTTTCCAAATCCGAAAAA 1733
DB |||||
QY 1681 GATTGGATGATTAATAATAAATTTGAAATGTTTCTTTTCCAAATCCGAAAAA 1733
DB |||||
```

RESULT 2

```
US-10-759-813-1
; Sequence 1, Application US/10759813
; Publication No. US20040139499A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-
; FILE REFERENCE: BBI465 US NA
; CURRENT APPLICATION NUMBER: US/10/759,813
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/909,566C
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219833
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1733
; TYPE: DNA
; ORGANISM: Euphorbia lagascae
US-10-759-813-1
```

```
Query Match 100.0%; Score 1733; DB 8; Length 1733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCATAAAAGGAAAATGGAGCAGAAAATCTCTCTTTCCGAGCATTTTAAAGTTTTCT 60
DB 1 GCATAAAAGGAAAATGGAGCAGAAAATCTCTCTTTCCGAGCATTTTAAAGTTTTCT 60
QY 61 GCTTGTGTTTAACTCTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGG 120
DB 61 GCTTGTGTTTAACTCTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGG 120
QY 121 GCCATGGAAGTTTCCATCATAGGTAATCTTCCCTCATTTATTACTCAGTCTTGATCTAGG 180
DB 121 GCCATGGAAGTTTCCATCATAGGTAATCTTCCCTCATTTATTACTCAGTCTTGATCTAGG 180
QY 181 CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGG 240
DB 181 CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGG 240
QY 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCTCAGCTGAAGCAGCAAGAGGTTATGAAAACTCA 300
```

[illegible]

RESULT 3									
US-10-686-947-153									
; Sequence 153, Application US/10686947									
; Publication No. US20040162420A1									
; GENERAL INFORMATION:									
; APPLICANT: Profigen Inc..									
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco									
; FILE REFERENCE: 79601									
; CURRENT APPLICATION NUMBER: US/10/686,947									
; CURRENT FILING DATE: 2003-10-16									
; PRIOR APPLICATION NUMBER: US 10/387346									
; PRIOR FILING DATE: 2003-03-12									
; NUMBER OF SEQ ID NOS: 298									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 153									
; LENGTH: 1538									
; TYPE: DNA									
; ORGANISM: NICOTIANATABACUM									
US-10-686-947-153									
Query Match 23.7%; Score 410; DB 8; Length 1538;									
Best Local Similarity 58.1%; Pred. No. 4.4e-91;									
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;									
Qy	113	CTCTCAGGGCCATGGAAGTTTCTCTATCATAGTAATCTTCCTCATATTATTA	CTCTCATCTTCT	172					
Db	129	CCACCAGGTCATGGAAACTACCAATCTAGGAAGTATGCTTCATATGGT-----TGGT	182						
Qy	173	GATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTATGGACCTGTTATGAGTCTT	232						
Db	183	GGACTACCAACACATGCTCTTAGAGATTTAGCCAAAATAATGGACCCGCTTATGCACCTT	242						
Qy	233	CAAAATGGGCCAAGTTTTCAGCTGTGTGCATTTCTTCAGCTGAAGCAGCCAAAGAGAGTTATG	292						
Db	243	CAATTAGGTGAAGTTTCTGCAGTTTGTGGTTACTTCTCTGATATGGCAAGAAGTACTA	302						
Qy	293	AAAACCTCAGCTGATGCGCTTTCGCCAAGCCCTATCGCTTTGGAGCGCACAGTTGTGTTT	352						
Db	303	AAAACCTCATGATCGCTTTCGCGCTTAGGCCCTAGCCCTTTTGGCCCCGGAGATTGCTGT	362						
Qy	353	TATAATCGGAAGATGCTCTTGTTCCTCATATGAGATCACTGGAGCGCATCAAGAA	412						
Db	363	TACAAATAGGTCTGATCTTCGCTTTTGGCCCTATGGCGATTATGGAGACAATGCGTAA	422						
Qy	413	ATTTGGATCTTGAATTTCTGAGTGCCAAAAGTTTCAATCCTCCAGGTTAATCCGAGAG	472						

Db 423 ATATGTGCTTGGAAAGTCTCAGTGCCCAAGAAATGTTCCGACATATAGCTCTATTAGGCGC 482
Qy 473 GAAGAAATGGAGGATGCCATCACATCTCTCGGTTCGAAAGCGGATCTCCGGTCAATATT 532
Db 483 GATGAAGTCTTCGTCTCTTAATTTTATCCGGTCATCTTCTGGTGAGCCCTGTTAATATT 542
Qy 533 ACAAGATCATTTATGGCATTATATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGATAGATCAGCGTTTGGGCAAGTA 602
Qy 590 TGTAAACAAAGAAAGATGCTGAGTGTGCCGATGCAGTCAATGAGGCGAGCAGAGT 649
Db 603 TTCAAGSAGCAGACAAATTTATACAATAATTAAGAAAGTTATATCTTTAGCAGGAGG 662
Qy 650 TTTGCGCCGAGCGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGGAGCTCAG 709
Db 663 TTTGATGTGGCTGACATATCCCTTCATCAAGTCTCTTCATGTCTCAGTGGATGAAG 722
Qy 710 TCAAAACCCAGCGGTTTGCATCAGAGATGACGATATACCTTGAAGAGATCTTAAATGAA 769
Db 723 GGTAAATATTGATGCAACACATAGGTAGATGCTATTGTTGAGAATGTCATCAACGAG 782
Qy 770 CAC-----AAAGCCATAGCCTTTTGAAGCGGATTAATATG 808
Db 783 CACAAGAAAAATCTTGCAATTTGGGAAATTAATGAGCGTTAGGAGGTGAAGATTTAATT 842
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTGCCAGTGACAAAAACGAAAGC 868
Db 843 GATGTTCTTAAACTTATGATGATGGAGCCTTCAATTTCTCTATCACCACGACAC 902
Qy 869 ATCAAGCATCCGTTTGGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
Db 903 ATCAAGCTATAATCTTTGACATGTTTGTGCTGGAACAGAGACTTCATCGTCAACAATT 962
Qy 929 GNAATGGTAAAGCGAGCTGATGAATAATCAACTGAACCTAAGAAAGCACAAGAA 988
Db 963 GTGTGGGCTATGGTGAATGGTGAATAATCCAATGATTTTGGAAAGCTCAACAGAA 1022
Qy 989 GTTAGCAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1023 GTAAGAGATGATTTAGAGAAAGAAAGAACTTTTGTATGAATAATGATGGAGGAGTAAAC 1082
Qy 1049 TTCCTCAAGTTAGTGGTTAAAGAAACTCTAAGATTAATCTCCCGGT---GTCCTGATT 1105
Db 1083 TATCTAAAGTTAGTCAATTAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1142
Qy 1106 CCGAGGAGGTAGAGAAACACGAAATTTGATGATGAAATTCATCCGAACTCGA 1165
Db 1143 CNAAGAAATGTAGGAAAGAGACAAATATAACGGCTACATTTCTGTGAAGACCAA 1202
Qy 1166 ATTTGTTGAAATGCTTTGGGCGATAGGAAGAGATCCCTAATCTTGGTGGAACTCGGAAAG 1225
Db 1203 GTCATGTTAATGTTTGGGCAATGGGAGAGATCCAAAATAATGGGATGATCGAGAACT 1262
Qy 1226 TTTAACCCGAAAGGTTTAAAGATTGTGCAATTGATTAAGAGGAGCAGCATTTGAACTG 1285
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGTTGTAATAATTTTGAATAT 1322
Qy 1286 GTACCAATTTGGTGCAGGAAAGAAATATGCTCGGCATTTACTTCAGCTATTACCAATTTG 1345
Db 1323 CTTCCATTTGGTGGGAAAGGAGGATTTGTCCAGGATTTCTGTTGTTTACCTAATGCT 1382
Qy 1346 GAGTATGCTATTATAAATCTATTATATCAATTTTAAATGGGAACCTGGCCGATGGAATTACA 1405
Db 1383 TATTTGCCATTGGCTCAATTAATCTTTATCACTTTTGAATGGGAATCCCACTGGAATCAA 1442
Qy 1406 CCTCAAAACATTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAGAAATATAGATCTT 1465
Db 1443 CCAAGCGACTTGGACTTGAATCTGAGTTGGTGGAGTAACCTGCGCTAGAAAAAAGTACCTT 1502
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1503 TACTTGGTTCGACTCTCTATCAA 1526

RESULT 4

US-10-943-507-153
; Sequence 153, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
; US-10-943-507-153

Query Match 23.7%; Score 410; DB 10; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CCTCCAGGCGCATGGAAGTTTCTCATCATAGTAACTCTCTCATTTTACTTACTCATTCT 172
Db 129 CCACAGGTCTCATGGAACCTACCAATCTAGGAAGTATGCTTCATATGGT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTGTAGAGCCTTGGCTCAAAATTTATGGACCTGTTTATGAGTCTT 232
Db 183 GGACTACACACCATGTCCTTAGAGATTTAGCCAAAAATATGGACCGCTTATGCACCTT 242
Qy 233 CAATGTCGCAAGTTTTCAGCTGTTGTCATTTCTTCAGCTGAAGCAGCACAAGAGGTTATG 292
Db 243 CAATTAGTGAAAGTTTCTGCAGTTGTGTGTTACTTCTCTGATATGCGCAAAAGAGTACTA 302
Qy 293 AAAACTAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGAGCGCACAGATTGTGTTT 352
Db 303 AAACTCATGACATCGCTTTCGGTCTAGGCTAGCCCTTTTGGCCCGGAGATTGTCTGT 362
Qy 353 TATAATCGGAAGATGCTTGTGTTGCTTTCATATGGAGATCACTGGAGCGAGATGAAGAAA 412
Db 363 TACAATAGTCTGATCTTGGCTTTTGGCCCTATGGCGATTAATGGAGACAAATGCGTAAA 422
Qy 413 ATTTGGATCTTGAATTTCTGAGTCCCAAAAAGTTTCAATCTCCAGGTTAATCCGAGAG 472
Db 423 ATATGTCTTGGAAAGTGTCTCAGTCCCAAGAAATGTTCCGACATATAGCTCTATTAGGCC 482
Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTCGAAAGCCGATCTCCGCTCAATATT 532
Db 483 GATGAAGTCTTCTGCTCTCTTAATTTTATCCGGTCACTCTTCTGGTGAGCTGTTAATATT 542
Qy 533 ACAAGATCATTTATGGCAATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602
Qy 590 TGTAAAGCAAAAGAAAGATTTGCTGAGTGTTCGGATGAGTCAATGAGGACGAGCACT 649
Db 603 TTCAAGGAGCAGACAAAATTTATACAACCTAAATTAAGAAAGTTTATCTCTTTAGCAGG 662
Qy 650 TTTGCAACCGCAGACGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGAGGCTGAG 709

```
Db 663 TTTGATGTGGCTGACATATTTCCCTTCATACAAGTCTCTTCATGTGCTCAGTGGAAATGAAG 722
Qy 710 TCAAAACCCAGGCGTTTGATCAGGAGATTCAGCATATATCTTGAAGAGATTTCTTAATGAA 769
Db 723 GGTAAGATTATGAATGCACACCATAGGATAGTCTATTTGTTGAGAAATGTCATCAACGAG 782
Qy 770 CAC-----AAAGCCAAATAAGCCTTTTGAAGCGATAACTTAAATG 808
Db 783 CACAAGAAAAATCTTCCAAATTCGGAAAACTATATGGAGCGTTAGGAGGTGAAGTTTAAT 842
Qy 809 GATGTTCTATGTAATCTTCAAAAAATGGAAGAGTTTCCAGTGCCAGTGACAAACGAAAGC 868
Db 843 GATGTTCTTCTAAACCTTATGAATGATGGAGGCGCTTCAATTTCTTATCAACCAACGACAAC 902
Qy 869 ATCAAGACATCCGTTTTCGAAATGTTTACTGCGGGAGCGGAACAACTTCGAAAGCTACA 928
Db 903 ATCAAGACATTAATCTTTGACATGTTTGTCTGGAACAGAGACTTCATCGTCAACAAT 962
Qy 929 GAATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACTAAGAAAAACACAAGAA 988
Db 963 GTGTGGCTATGTTGGAATGTTGAAAAATCCAACTGATTTTCCGAAAGCTCAAGCAGAA 1022
Qy 989 GTTAGCAAGATTTTGGTGAATGGAAGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1023 GTAAGAGATGCTTTAGAGAAAAAGAAACTTTTGTATGAAAAATGATGTGGAGGAGCTAAAC 1082
Qy 1049 TTTCTCAAGTTAGTTGTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCATTGATT 1105
Db 1083 TATCTAAAGTTAGTCTATTAAGAAACTCTAAGACTTTCAATCCACCGGTTCCACTTTTGCTC 1142
Qy 1106 CCGAGGAGGTAGAGAAACAAACAGAAATGATGATATGAAATTCATCCGAACACTCGA 1165
Db 1143 CCAAGAGATGTAGGAGAGACAAATATAACGGCTACACTATTTCTGTAAAGCCAAA 1202
Qy 1166 ATGTTGTGAATGCTTTGGCGATAGGAAGAGATTCCTTAATCTTGTGCGAACTCGAAAG 1225
Db 1203 GTCATGTTTAATGTTTGGGCATTTGGGAGAGATCCAAATAATTTGGGATGATCAGAACT 1262
Qy 1226 TTTAACCCAGAAAGGTTTAAAGATGTGCAATGATTATAAGGAGACATTTGAACTG 1285
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGTTAAATTTTGAATAT 1322
Qy 1286 GTACCAATTTGGTGAGGAAAAAGAAATATGCTCTGGCATTACTTCAGCTATTACCAATTTG 1345
Db 1323 CTTCATTTGGTGGTGAAGAGGATTTGTCCAGGATTTGTTGGTTAGCTAATGCT 1382
Qy 1346 GAGTATGCTATTAATCTATTATATCAATTTTAAATTTGGAACTGGCCGATGGAATFACA 1405
Db 1383 TATTTGGCAATGGCTCAATTAATCTTTTATCACTTTGATTTGGAACTCCCACTGGAATCAA 1442
Qy 1406 CCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1443 CCAAGCGACTTGGACTGACTGAGTTGGTGGAGTAACATGCGGCTAGAAAAAGTGAACCTT 1502
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1503 TACTTGGTTGCGACTCCTTATCAA 1526
```

RESULT 5

```
US-10-934-944-153
; Sequence 153, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
```

```
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; US-10-934-944-153
```

```
Query Match 23.7%; Score 410; DB 11; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CTCACAGGCGCCATGGAAGTTTCCTATCATAGTAGTAATCTTCTCTCATTTATTACTCACTTCT 172
Db 129 CCACCAGGTCATGGAACACTACCAATAGGAAGTATGCTTCATATGCT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTGTAGAGCTTTGGCTCAAAATTTATGGACCTGTTTATGATCTTT 232
Db 183 GGACTACCAACACCATGCTCTTAGAGATTTAGCCAAAAAATATGGACCGCTTATGCACCTT 242
Qy 233 CAAATTTGCCAAGTTTTCAGCTGTTGCTTCTTCCAGCTGAGAGCCCAAGAGGTTATG 292
Db 243 CAATTAGGTGAAGTTTCTGCAAGTTTGTGCTTCTCTCTGATATGGCAAAAGTACTFA 302
Qy 293 AAAACTCAGGCTGATGCTCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTT 352
Db 303 AAAACTCATGACATGCTTTTCGCGCTAGGCTAGGCTTTTGGCCCCGGAGATTTGCTCT 362
Qy 353 TATAATCCGAAGATGCTGTTGTTGCTTCAATATGAGATCACTGGAGCGAGATGAAGAAA 412
Db 363 TACAATAGGTCTGATCTTGGCGTTTTCGCTTATGGCGATTTATGGAGACAAAATGCGTAAA 422
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCTTCAATCTTCCAGGTTTCAATCTTCCAGGTTTATCCGAGAG 472
Db 423 ATATGTGCTTTGGAAGTGTCTAGTGCCTTCAAGATTTTCGACATATAGCTCTATTTAGGCGC 482
Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCCGTTCCGAAAGCCGAGATCTCCGCTCAATATT 532
Db 483 GATGAAGTTCTTCGCTCTCTTAATTTTATCCGCTCATCTTCTGTTGAGCGCTGTTAATATT 542
Qy 533 ACAAGATCATTTATGGCATTTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATGAGCGTTTGGGCAAGTA 602
Qy 590 TGTAAGCAAAAAGAAAGATTTGCTGAGTGTGCGGATGCTCAATGAGGACGACGAGT 649
Db 603 TTCAAGGAGCAAGACAAATTTATCAACTAATTAAGAAGTTTATCTCTTACGAGAGGG 662
Qy 650 TTTGSCACCGCAGAGCGCTTTTCCGACGTGGAATTTACTACTATATCATTTGGAGCTGAG 709
Db 663 TTTGATGTGGCTGACATATTCCTTTCACAAGTCTCTTCATGCTCTCAGTGAATGAAG 722
Qy 710 TCAAAAACCCAGGCGTTTTCGATCAGGAGATTTGACGATATATCTTGAAGAGATTTCTTAATGAA 769
Db 723 GGTAAGATTATGAATGCACACATAAGGATAGTGTATTTGTTGAGAAATGTCTATCAACGAG 782
```

```
QY 770 CAC-----AAAGCCAAATAGCCTTTTGAAGCGGATAACTTAATG 808
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 CACAAGAAAAATCTTTGCAATTCGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTAAAT 842
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 809 GATGTTCTATTGAATCTTCARAAAATGGAACGTTCCAGTCCAGTGACAAACGAAAGC 868
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 GATGTTCTTCTAAAATCTTATGAATGATGGAGCGCTTCAATTTCCCTATCACCAACGACAA 902
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 869 ATCAAGACATCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 ATCAAGCTATAATCTTTGACATGTTTCTGCTGGAACAGAGACTTCATCGTCAACAAT 962
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 929 GAATGGTAAATGGCAGAGCTGATGAAAATCCAATCAACTGAATGAAGAAACACAAAGAA 988
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 GTGTGGCTATGCTGGAATGCGAAAAATCCAATGTATTTGCGAAAGCTCAAGCAGAA 1022
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 989 GTTAGCAGATGTTTGGTGAATGCGGAAGGTTGATGAATCAAGATTTCAATGATTTGAA 1048
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1023 GTAAGAGATGCAATTTAGAGAAAAGAACTTTTGATGAAAATGATGTGGAGGAGCTAAAC 1082
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1049 TTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCCGTT---GTCTTGATT 1105
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1083 TATCTAAAGTTAGTCATTAAGAAACTCTAAGACTTCATCCACCGTTCCACTTTTGTCT 1142
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1106 CCGAGGAGGTAGAGAAACACAGAAATGTGATGATGAAATTCATCCGAACTCGA 1165
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1143 CCAAGAGATGTAGGGAAGAGACAAATATAAACGGCTACACTATTTCTGTAAAGACCAA 1202
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1166 ATTTGTGTAATGCTTGGCGGATAGGAAGAGATCTTAATACCTTGTGCGAACCTGGAAG 1225
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1203 GTCATGTTAATGTTTGGGCATTTGGGAAGATGCAAAATATGGATGATGCAAACT 1262
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1226 TTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAATAAGGAGCAGACATTTGAAC 1285
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTCTTAAGGATTTTGTGTTAAATTTTGAATAT 1322
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1286 GTACATTTGGTGGAGGAAAAGATATGTCCTGGCATTTACTTCAGCTATTACCAATTTG 1345
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1323 CTTCATTTGGTGGTGAAGGAGGATTTGTCCAGGATTTCTGTTGGTTAGCTAATGCT 1382
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATGGGAACCTGGCGATGGAATTACA 1405
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1383 TATTTGGCATGGCTCAATTACTTTATCACTTTTGAATGGAACTCCCACTGGATCAA 1442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1406 CTTCAAACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATGATCTT 1465
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1443 CCAAGCGACTTGGACTTGACTGAGTTGGTGGAGTAACTGCGGCTAGAAAAAGTGACCTT 1502
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1503 TACTTGGTTGGCACTCCTTTATCAA 1526
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 6

```
US-11-116-881A-162
; Sequence 162, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/565,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/565,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/566,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
```

```
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 162
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-11-116-881A-162
```

```
Query Match 23.7%; Score 410; DB 16; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTTATCATAGTAACTCTCCTCATTTTACTACTACTTCT 172
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 CCACCAGGTCCATCGAACTACCAATCTAGGAAGTATGCTTCATATGGT-----TGGT 182
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 173 GATCTAGGCCATGAACGTTTGTAGAGCCTTGCTCAAAATTTATGGACCTGTTTATGAGTCTT 232
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 GGACTACACACCATGCTTAGAGATTTAGCCAAAATATGGACCGCTTATGACCTT 242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 CAAATTGGCCAAAGTTTCAGCTGTTGTCAATTTCTTCAGCTGAAGCAGCAAGAGAGTTATG 292
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CAATTAGGTGAAGTTTCTGCAGTTGTGGTTACTTCTCTGTATATGGCAAAAGAGACTACTA 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 AAAACTCAGGTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACAGATTCGTGTTT 352
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AAAACTCATGACATGCGCTTTCGCGTCTAGGCGCTAGCCCTTTTGGCCCCGGAGATTTCTGT 362
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 TATAATCGAAAGATGCTCTGTTTGTCTTATATGGAGATCACTGGAGGCGAGATGAAGAAA 412
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 TACAATAGTCTGATCTTGCGTTTGGCCCTATGGCGATTTATGGAGACAAATGCGTAAA 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 ATTTGGATACCTTGAATTTCTGAGTGCCAAAAGATTTCAATCTCCAGGTTAATCCGAGAG 472
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 ATATGTGCTTTGGAAAGTCTCAGTGCCAAAGAAATGTTTCGGACATATAGCTCTATTAGGCGC 482
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTTCGAAAGCCGGATCTCCGGTCAATATT 532
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GATGAAGTTCTTCGTCTCTTAATTTTATCCGTCATCTTCTGGTGAGCGCTGTTAATATT 542
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 ACAAGATCATTTATGGCAATTATAATTTGATCATGATAAGAACATCCGTTGGTAA---T 589
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 ACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 590 TGTAAGCAAAAGAAAGATTTGCTGAGTGTTCGCGATGCAATGAGCAGCAGCAGAGT 649
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 TTCAAGGAGCAAGACAAAATTTATACAACTAATTAAGAAAGTTATATCTCTTAGCAGAGGG 662
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 650 TTTGCGACCGCAGAGCGCTTTTCGCGAGTGGAAATTTACTATATATCATTCGAGCTGAG 709
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 TTTGATGTGGCTGACATATTCCTTTCATACAGTCTCTTTCATGTCTCAGTGAATGAAG 722
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 710 TCAAAACCCAGCGGTTTTCATCAGGAGATTTGACGATATATCTTTGAAGAGATTTCTTAATGA 769
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 GGTAAGATTATGAATGCACACACCAATGAAGGTAGATGCTATTGTTGAGAAATGTCATCAACGAG 782
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 770 CAC-----AAAGCCAAATAGCCTTTTGAAGCGGATAACTTAATG 808
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 CACAAGAAAAATCTTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTAAAT 842
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACGTTTCCAGTCCAGTGCACAAACGAAAGC 868
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 843 GATGTTCTTCTTAAACTTATGAATGATGGAGGCTTCAATTTCTCTATCAACAACGACAAC 902
QY 869 ATCAAGACATCCGTTTTCGAAAGTTTACTGCGGGAGCGAACAACACTTCGAAAGCTACA 928
Db 903 ATCAAGACTATAATCTTTGACATGTTTGTCTGCTGGAAACAGAGACTTCATCGTCAACAAT 962
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAATAATCCAACTGAACTAAGAAAGACACAAGAGAA 988
Db 963 GTGTGGCTATGTGGAAATGGTGAATAATCCAACTGATTTGCGAAAGCTCAAGCAGAA 1022
QY 989 GTTAGACAAGATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1023 GTAAGAGATGCAATTTAGAGAAAAGAACTTTTGTGATGAAAATGATGGAGGAGCTAAAC 1082
QY 1049 TTCTTCAAGTTAGTGTGTAAGAACTCTAAGATTACATCTCCGTT---GTCCTGATT 1105
Db 1083 TATCTAAGTTAGTCAATTAAGAAACTCTAAGACTTCATCCACCGTTCCACTTTTGTCTC 1142
QY 1106 CCGAGGAGTGTAGAGAAACAAACAGAAATGTGATGATGAAATTCATCCGAACACTCGA 1165
Db 1143 CCAAGAGATGTAGGAGAGACAAATATAAACGGCTACACTATTCTGTAAAGACAAA 1202
QY 1166 ATGTTGTGAATCTTGGGCGATAGGAAGAGATCTTAATACTTTGTTGGAACCTGGAAAG 1225
Db 1203 GTCATGGTTAAATGTTGGGCAATGGGAAGAGATCCAAATATTGGGATGATGCAGAACT 1262
QY 1226 TTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAAGAGGAGACATTTGAACATG 1285
Db 1263 TTAAAGCCAGAGAGATTTGAGCAGTGTCTTAAGGATTTTGTGGTAAATAATTTGAAAT 1322
QY 1286 GTACCAATTTGGTGCAGAAAAGAAATATGTCCTGGCATTACTTCAGCTATTACCAATTTG 1345
Db 1323 CTTCATTTGGTGTGAGGAGGATTTGTCCAGGATTTGTTGGTTAGTAAAGTCT 1382
QY 1346 GAGTATGTCATTAATAATCTATTATATCAATTTTAAATGGGAATGCGCGGATGGAATPACA 1405
Db 1383 TATTTGCCATTGCTCAATTAATTTATCACTTTTGAATGGGAATCCCCACTGGAATCAAA 1442
QY 1406 CCTCAACACTTGATATGACTGAGCTATTGGCGGTGCTCAGGAAAAAATAGATCTT 1465
Db 1443 CCAAGCGACTGGACTTGACTGATGTTGGTGGAGTAACTGCCGCTAGAAAAAGTGACCTT 1502
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1503 TACTTGGTTGGACTCTCTATCAA 1526

RESULT 7

US-10-686-947-209
; Sequence 209, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 209
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-209

Query Match 23.1%; Score 400.4; DB 8; Length 1673;

Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCCATGGAAGTTTCTCTATCATAGGTAATCTTCCTCATTTATTACTCTCTTCT 172

Db 135 CCACCAAGTCCATGGAAACTACCAATACTAGGAAGTATGCTTCATATGGT-----TGGT 188
QY 173 GATCTAGGCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGAACCTGTTATGAGTCTT 232
Db 189 GGACTACCAACACCATGCTCTTAGAGATTTAGCCAAAAAATATGGACCACTTATGCACTT 248
QY 233 CAAATTTGGCCAAAGTTTTCAGCTGTTGTCATTTCTTCACTGTAAGCAGCAAGAGGTTATG 292
Db 249 CAAATAGGTGAAGTTTCTCGGTTTGGTTTACTTCTCTGATACGGCAAAAGAGATTTA 308
QY 293 AAAAATCAGGCTGATGCCCTTTCGCCCAACGCCCTTATCGTCTTGGACGCACACAGATTGTTT 352
Db 309 AAAAATCAGATGCAATCGCTTTTGGCTTAGGCCCTAGCTTTTGGCCCCGGAGATTGCTGT 368
QY 353 TATATTCGAAAGATGCTGTTGTTTGGTTTCAATGAGAGATCACTGGAGGAGATGAAAGAAA 412
Db 369 TACAATAGGTCTGATCTAGCCTTTTGGCCCTATAGCGACTATTGGAGACAATAATGCGTAAA 428
QY 413 ATTTGGATACTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTTAATCCGAGAG 472
Db 429 ATATGTTCTTGGAGTGTCTCAGTGCACAAAGATGTTTCGACATTTAGCTCTTATTAGCGCG 488
QY 473 GAAGAAATGGAGGATGCCATCACATTCCTCCGTTTCGAAAGCCGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGTTCTCATTAATTTTATCCGCTCATCTTCTGGTGAACCTATTATGTT 548
QY 533 ACAAGATCATTTATGGCAATTAATTTTCGATCATGATGAAGAACATCCGTTGGTAA---T 589
Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGATAGATCAGCGTTTGGGCAAGTG 608
QY 590 TGTAGCAAAAAGAAAGATTGCTGAGTGTTCGCGATGCACTCAATGAGCAGCAGCAGT 649
Db 609 TTCAAAGAGCAAGACAAATTTATACAATTAATAAGAAAGTGATCTCTTAGCAGGAGG 668
QY 650 TTTGGCACGCGACAGCGCTTTTCCGACGTGGAAATTACTTCACTATATCATTTGAGCTGAG 709
Db 669 TTTGATGTCGTGACATATTCCTTCACTGAAGTTTCTTCATGTCGTGAGTGAATGAAG 728
QY 710 TCAAAACCCAGCGTTTGCATCAGGAGATGACGATATACCTTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCATTAAGGTAGATGCCATTTGTGAGAAATGTCATCAATG 788
QY 770 CAC-----AAAGCCATAAGCCTTTTGAAGCGGATTAACCTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAATAGGAGCGTTTAGGAGGTGAAGATTTAAT 848
QY 809 GATGTTTCTTATGAATCTTCAAAAAAATGAAAAAGTTTCCAGTGCCAGTGACAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGCGCTTCAATTTCTTATCACCACGACAAC 908
QY 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGACTACA 928
Db 909 ATCAAAGCTATAAATTTTGAATGTTTGTCTGCGGGACAGAGACTTCATCGTCAACAAAT 968
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAATGAAGAAAAAGCACAAGAGAA 988
Db 969 GTGTGGCTATGTTAGAAATGGTGAANAATCCAGCCGTTATTCGCAAGAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGAAAAAGTTGATGAATCAAGATTTCTCATGTTGAAA 1048
Db 1029 GTAAGAGAGCATTATAGAGAAAAGAACTTTCGATGAATAATGATGTGGAGAGCTAAAC 1088
QY 1049 TTTCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCGTT---GTCCTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTCACTCCACCGGTTCCACTTTTGTCTC 1148
QY 1106 CCGAGGAGTGTAGAGAAACACACGAATTTGATGATATGAATTAATTCATCCGAACACTCGA 1165
Db 1149 CCAAGAGAAATGTAGGAAAGAGACAAATATAAACCGCTACACTTTCTCTGTAAAGACCAA 1208
QY 1166 ATTTGTTGAAATGCTTGGCGGATAGGAAGATCTTAATCTTGTGCGGAACCTGGAAG 1225

Db 1209 GTCATGGTTAAATGTTTGGGCTTTGGGAGAGAGATCCAAATAATTTGGAATGACCCAGAAACT 1268
Qy 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGAGGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTTGGTGAATAATTTTGAATAT 1328
Qy 1286 GTACCAATTTGGTGAGGAAAAAGAAATATGTCCTGGCAATTAATTCAGCTATTAACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGGGAAGGAGGATTTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATTTGGAATCGCCGATGCAATTAACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTAATCACTATATCACTTCGATTTGGAATCTCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAAAACATTTGATATGACTGAAGCTATTTGGCGGTGCTCTCAGGAAAAAATATAGATCTT 1465
Db 1449 CCAAGCGACTTTGACTTGACTGAGTTGGTTGGAGTAACTGCCGCTAGAAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCCTTATCAA 1532

RESULT 8

US-10-686-947-211
; Sequence 211, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 211
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-211

Query Match 23.1%; Score 400.4; DB 8; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
Qy 113 CTTCCAGGCGCCATGGAAGTTTCTCATCATAGTGAATCTTCTCATTTTATTACTCACTTCT 172
Db 135 CCACAGGTCCATGGAACTACCAATACTAGGAAGTATGCTTCATATGTT-----TGGT 188
Qy 173 GATCTAGGCCATGAAGTTTATAGAGCTTTGGCTCAAAATTTATGGACCTGTTATAGTCTT 232
Db 189 GGACTACCAACACCATGCTCTTAGAGATTTAGCCAAAAAATAATGGACCACTTATGCACCTT 248
Qy 233 CAAATTTGGCCAAATTTTCACTGTTGTCATTTCTTCACTGGAAGCAGCCAAAGAGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTGCGTTTGGTTACTTCTCTGATACCGCAAAAGAAATGATTA 308
Qy 293 AAAACTCAGGCTGATGCTCTTCCGCAACGCCCTATGCTCTTGGAGCGACAGATTGTTT 352
Db 309 AAACCTCATGACATCGCTTTTGGCTTAGCCCTTAGCCCTTTGGCCCGGAGATTGCTGT 368
Qy 353 TATAATCGGAAAGATGTCTTGTGTTGCTTCATATGAGATCACTGGAGCGAGATGAAGAAA 412
Db 369 TACAAATAGTCTGATCTAGCTGCTTTTGGCTTAGCCCTTAGCCCTTTGGCCCGGAGATTGCTGT 428
Qy 413 ATTTGATTAATTTCTGATGCGCAAAAGTTTCAATCTTCAAGTTTATTCAGAG 472
Db 429 ATATGTGCTTGGAAAGTGTCTAGTGCAGAAAGATTTTCGACATTTTAGCTCTATTAGGCGG 488
Qy 473 GAAGAAATGGAGATGCCATCACTTCTCCGTTCGAAAGCCGGATCTCCGCTCAATATT 532

Db 489 AATGAAGTTCTTGTCTCAATTAATTTTATCGGTCACTCTTCTGGTGAACCTATTAAATGTT 548
Qy 533 ACAAGATCAATTTATGGCAATTAATTTTGCATCATGATAAAGAACATCCGTTGGTAA---T 589
Db 549 ACGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTTCGCCGATGCACTCAATGAGGCGAGCAGAGT 649
Db 609 TTCAAAGAGCAAGCAAAATTTATACAACATAAATAAAGAAAGTGATACTCTTAGCAGAGGG 668
Qy 650 TTTGSCACCGCAGAGCGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTATGGAGCTGAG 709
Db 669 TTTGATGTGGCTGACATATTTCCCTTCACTGAGATTTCTTCACTGTCTCAGTGAATGAAG 728
Qy 710 TCAAAACCCAGCGCTTTGCATCAGGAGATTGACGATATATCTTGAAGAGATTCTTTAATGAA 769
Db 729 GGTAAGATTATGAATGCACACACATGAAGGTAGATGCCATTTGTTGAGAATGTATCAATGAG 788
Qy 770 CAC-----AAAGCCCAATAGCCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGAAAACTTCCAGTGCCAGTGACAAAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGCCCTTCAATTTCTCTATCACCACGACAAAC 908
Qy 869 ATCAAGCATCCGTTTTCGCAAAATGTTTACTGCGGGAGCGAAACAACTTTCGAAAGCTACA 928
Db 909 ATCAAGGCATAATTTTTCGACATGTTTCTGCTCGGGACAGAGACTTTCATCGTCAACAAAT 968
Qy 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCACAAGAA 988
Db 969 GTGTGGGCTATGGTAGAATAATGGTAAAAATCCAGCCGTATTCGCGAAAGCTCAACAGAA 1028
Qy 989 GTTAGACAAGTATTGGTGAATGGAAAAAGTTGATGAATCAAGATTTTCATGATTGAAA 1048
Db 1029 GTAGAGAGCATTTAGAGGAAAGAAACTTTCGATGAAATATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TCTTCAAGTTAGTGGTTAAAGAACTCTTAAGATTATACCTCCCGGTT---GTCCTTGAAT 1105
Db 1089 TACCTAAAGTTAGTAATAAAAAAGAACTCTAAGACTTTCACCCCGTTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGTGTAGAGAAACACAGAACTTGTGATGATGAAATTCATCCGAAACACTCGA 1165
Db 1149 CCAAGAGAAATGTAGGAAAGAGACAAATATAACCGCTACACTTTCTCTGTAAGACCAA 1208
Qy 1166 ATTGTTGTGAATGCTTTGGCGATAGGAAAGAGATCCTAATACCTTGGTTCGGAACCTGGAAG 1225
Db 1209 GTCATGGTTAATGTTTGGGCTTTTGGGAAGAGATCCAAATAATTTGGAATGACGCAAACT 1268
Qy 1226 TTTAACCAGAAAGGTTTAAAGATTGTCGAATTTGATTAATAAGGAGCAGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTTGTGTAATAATTTTGAATAT 1328
Qy 1286 GTACCATTTGTCAGGAAAAAGAAATATGCTCTGCAATTTACTTCACTGATTTTACCAATTTG 1345
Db 1329 CTTCCATTTGTTGGGGAAGGAGGATTTGTCTCGGATTTGCTTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATTCGGAACTGGCCGATGGAATTA 1405
Db 1389 TATTGGCCATTTGGCTCAATTAATCACTATATCACTTCGATTTGGAACCTCCCTGCTGAATCGAA 1448
Qy 1406 CCTCAACACTTGTATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATAGATCTT 1465
Db 1449 CCAAGCGACTTGGACTTGAAGTTGAGTGGTTGGTGGAGTAACTCGCGCTAGAAAAAGTACCTT 1508
Qy 1466 AAGTTGATTCCTATTTCATATCAA 1489
Db 1509 TACTTGGTTGCACTCCTTATCAA 1532

; Sequence 231, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 231
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-231

Query Match 23.1%; Score 400.4; DB 8; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGCGCATCGAAGTTTCTATCATAGGTAACTTCCTCATTTTACTCACTTCT 172
DB 135 CCACCAAGTCCATGGAACCTACCAATACTAGGAAGTATGCTTCATATGGT-----TGGT 188
QY 173 GATCTAGGCGCATCAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTGTATGACTCTT 232
DB 189 GGNACTACCAACCATGCTCTTAGAGATTTTAGCCAAAATATAGGACCACTTATGCACTT 248
QY 233 CAAATTTGGCCAAAGTTTTCAGCTGTGTGTCATTTCTCAGCTGAAGCAGCCAAAGAGGTTATG 292
DB 249 CAATTAGGTGAAGTTTTCGGGTTGGTTTACTTCTCCTGATAGCGCAAAAGATTTA 308
QY 293 AAAACTCAGCTGATGCTTGGCCAAAGCCCTATGCTCTTGAACGACAGATTTGTTTT 352
DB 309 AAAACTCATGACATCGCTTTTGGCTTAGGCTTAGGCTTTTGGCCCGGAGATTTGCTGT 368
QY 353 TATAATCGGAAGATGCTTTGTTGCTTCATATGGAGATCACTGGAGGAGATGAAGAA 412
DB 369 TACAATAGGTCTGATAGCTTTTGGCCCTATGCGGCACTATTGGAGACAAATGCGTAAA 428
QY 413 ATTTGGATCTTGAATTTCTGAGTGCACAAAAGTTTCAATCTCCTCAGGTTAATCCGAGAG 472
DB 429 ATATGTGCTTGAAGTGTCTAGTGCCAAAGATGTTTCGGACATTTAGCTCTATTAGCGG 488
QY 473 GAAGAAATCGAGATGCCATCACAATTCCTCGTTGAAAGCCGAGTCTCCGGTCAATATT 532
DB 489 AATGAAGTTCTTTCGCTCATTAATTTATCCGTCATCTTCTGTTGAACCTATTAAATGTT 548
QY 533 ACAAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCCGTTGGTAA---T 589
DB 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGATAGATCAGCGTTTGGCAAGTG 608
QY 590 TGTAAAGCAAAAAGAAAGATTGCTGAGTGTGCGGATGCACTCAATGAGCAGCGACGAGT 649
DB 609 TTCAAAGAGCAAGACAAATTTATACAACCTAAATTAAGAAAGTGATCTCTTAGCAGGAG 668
QY 650 TTTGGCAGCGAGACGCTTTTCGAGCTGGAATTTACTTCACTATATCAATTTGAGCTGAG 709
DB 669 TTTGATGTGCGCTGACATATTCCTTCCCTTCACTGAAGTTTCTTCATGTGCTCAGTGGAA 728
QY 710 TCAAAACCCAGCGTTTTCATCAGGAGATTGAGATATATCTTGAAGAGATTTCTTAATGAA 769
DB 729 GTGATGATTATGAATGCACACCAATAGGTAGTAGCCATTTGTTGAGATGTCATCAATGAG 788
QY 770 CAC-----AAAGCCAATAAGCCTTTTGAAGCGGATAACTTTAATG 808
DB 789 CACAAGAAAAATCTTCCAAATTTGGAAACTAATAGGAGCGTTAGGAGTGAAGTTTAAT 848
QY 809 GATGTTCTAATGATCTTCAAAAAATGAAACGTTTCCAGTGCCAGTGACAAACGAAAGC 868
DB 849 GATGTTCTTAAAGACTTATGAATGATGAGGCGCTTCAATTTCTTCAATCAACCAACGAC 908

QY 869 ATCAAGCATCCGTTTGCAAATGTTTACTGCCGGGAGCGAACAACATTCGAAAGCTACA 928
DB 909 ATCAAAAGCCATAATTTTGTGATGTTTGTGCGGGACAGAGACTTTCATCGTCAACAAT 968
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACTAAGAAAAAGCACAAAGAA 988
DB 969 GTGTGGCTATGCTAGAAATGGTGAANAATCCAGCGTATTCCGGAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGGAAAAGTTGATGAATCAAGATTTCAATGATTTGAAA 1048
DB 1029 GTAAGAGAAAGCATTTAGAGGAAAAGAACTTTTCGATGAAAAATGATGCGGAGGAGCTAA 1088
QY 1049 TTTCTTCAAGTGTAGTGGTTAAAGAACTCTAAGATTACATCTCTCCGGTT---GTTCTGAT 1105
DB 1089 TACCTAAAGTTAGTAAATAAAGAACTCTAAGACTTCAACCCGGTTCACACTTTTGTGCT 1148
QY 1106 CCGAGGAGCTAGAGAAAACAACAGAAATTGATGGATATGAAATTCATCCGAACACTCGA 1165
DB 1149 CCAAGAGAAATGTAGGGAAGAGACAAATATAAAGCGCTACACTATTCTGTAAAGACCAA 1208
QY 1166 ATGTTGTCAATGCTTGGCGGATAGGAAGAGATCTTAATACTTTGTCGGAACCTGGAAAG 1225
DB 1209 GTCATGGTTAATGTTTGGGCTTTGGGAGAGATCCAAAATATTGGAATGACGCGAAGCT 1268
QY 1226 TTTAAACCCAGAAAGTTTAAAGATTGTGCAATTGATTATAAAGGAGCAGCATTTGAACTG 1285
DB 1269 TTTATGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGGTAAATAATTTTGAATAT 1328
QY 1286 GTACCAATTTGGTCAGGAAAAAGAAATATGCTCTGCGCATTTACTTCAGCTATTACCAATTTG 1345
DB 1329 CTTCCATTTGGTGGCGAAGGAGGATTTGCTCTGGGATTTCTTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGCTAATTAATACTTATATATCAATTTTAAATTTGGAACTGCGCCGATGGAATTA 1405
DB 1389 TATTTGCCATTTGGCTCAATTTACTATATCACTTTGATTTGGAATCTCCTGCTGGAATCG 1448
QY 1406 CCTCAACACACTTGATATGATCAAGCTATTGCGGCTGCTCTCAGGAAAAAATAGATCTT 1465
DB 1449 CCAAGGACTTGGACTTGACTGAGTTGGTGGAGTAACATGCGCGTAGAAAAAGTGACCTT 1508
QY 1466 AAGTTGATTCCTATTCCATATCAAA 1489
DB 1509 TACTTGGTTGCGACTCTTATCAA 1532

RESULT 10
US-10-943-507-206
; Sequence 206, Application US/10943507
; Publication No. US20050132444A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 206
; LENGTH: 1673
; TYPE: DNA

```
; ORGANISM: NICOTIANATABACUM
US-10-943-507-206

Query Match      23.1%; Score 400.4; DB 10; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTATCATAGTAATCTTCCATTTACTTACTCTCT 172
D 135 CCACAGAGTCCATGGAACACTACCAATACTAGGAAGTATGCTTCATATGCT 188
QY 173 GATCTAGGCCATGAAGTTTTHAGAGCTTGGCTCAAAATTATGGACCTGTTATGAGTCT 232
D 189 GGACTACCAACACCATGTCTTAGAGATTTAGCCNAAAAATATGSAACCACTTATGCACCTT 248
QY 233 CAAATTGGCCAAAGTTTTCAGCTGTCTGTCATTCTTCAGCTGAAGCAGCAAGAGGTTATG 292
D 249 CAAATTAGGTGAAGTTTCTGCGGTTTGTGTTACTTCTCTGTATACGCGCAAGAAGTATTA 308
QY 293 AAAAATCTAGGCTGATGCTCTTCCGCCCAACGCCCTATCGTCTTGGACGCAAGATGTGTTT 352
D 309 AAAAATCTAGCATGCTCTTTCGCTTAGGCCCTAGCCCTTTTGGCCCCGGAGATTGCTGT 368
QY 353 TATATCGGAAGATGTCTTGTGTTTCTTTCATATGAGATCATCTGGAGCAGATGAAGAA 412
D 369 TACAAATAGGCTGATCTAGCCCTTTTGGCCCCCTATGGCGACTATTGGAGACAAATGCGTAA 428
QY 413 ATTTGGATACTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
D 429 ATATGTGCTTGGAGTGTCTAGTGCCAGAGATGTTTCGGACATTTAGCTCTATTAGGCGG 488
QY 473 GAAGAAATGGAGGATGCCATACATTCCTCCGTTCCGATCAAGAACATCCGTTGCTCAATATT 532
D 489 AATGAAGTCTTTCGCTCATTAATTTATCCGCTCATCTTCTGCTGGAACCTATTAAATGTT 548
QY 533 ACAAGATCATTTATGGCATTATATTCGATCATGATGAAGAACATCCGTTGGTAA---T 589
D 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGATGATGAGCGTTTGGGCAAGTG 608
QY 590 TGTAAAGCAAAAGAAAGATTGCTGATGTTGCGGATGCGAGTCAATGAGCGACGACGAGT 649
D 609 TTCAAGAGCAAGACAAATTTATACACTAATTAAGAAGTGATCTCTTAGCAGGAGG 668
QY 650 TTTGGCAGCGAGAGCTTTTCCGAGCTGGAATTTACTTCACTATATCATTTGGAGCTGAG 709
D 669 TTTGATGTGGCTGACATATTCCTTCACTGAAAGTTTCTTCACTGTCTCAGTGGAAATGAAG 728
QY 710 TCACAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAATGAA 769
D 729 GGTAAAGATTATGAATGCACACCATAGGTAGTATGCCATTTGTTGAGAAATGTCATCATGAG 788
QY 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATAACTTAATG 808
D 789 CACAGAAAAATCTTCCATTTGGGAAACTATATGGAGCGTTAGGAGGTGAAGATTTAAT 848
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACCGTTCCAGTCCAGTGCAGTGAACAAAGCC 868
D 849 GATGTTCTTAAAGACTTATGAATGATGAGGCGCTTCAATTTCTTATCACCACCAACGACAC 908
QY 869 ATCAAGCATCCGTTTGGCAATGTTTACTGCGGGGCGCAACCACTTCGAAAGCTACA 928
D 909 ATCAAGCTATTAATTTTGGACATGTTTGTGCGGGGACAGAGACTTTCATCGTCAACAATT 968
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAAGTAAAGAAAAAGCAAGAAGAA 988
D 969 GTGTGGCTATGTTAGAAATGGTGAATAATCCAGCGTATTTCGCGAAAGCTCAAGCAGAA 1028
QY 989 GTTAGCAAGATATTTGGTGAATATGGGAAAGTTTATGAATCAAGATTTCAATGATTGAAA 1048
D 1029 GTAAGAGAGACATTTAGAGGAAAAAGAACTTTTCGATGAAAAATGATGTGGAGGAGCTAAAC 1088
QY 1049 TTCTTCAAGTTAGTGTGTTAAAGAAACTCTAAGATTACATCTCCCGGTT---GTCCTGATT 1105
```

```
Db 1089 TACCTAAAGTTAGTAATAATAAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGGAGTGTAGAGAAACACACGAATTCATGATATGAAATTCATCCGAAACACTCGA 1165
D 1149 CCAAGAGAATCTAGGGAAGAGACAAATATAAAACGCTACACTATTCTCTGTTAAAGACCAAA 1208
QY 1166 ATTGTTGGAATGCTTGGCGGATAGGAGAGATCTTAATCTTGGTCCGAACCTCGAAAG 1225
D 1209 GTCATGGTTAATGTTTGGGCTTTGGGAAGAGATCCAAAATATTGGAATGACGACAGAACT 1268
QY 1286 TTTAACCCAGAAAGGTTTAAAGATTGTGCAATTGATTATAAAGGACGACACATTTGCAACTG 1285
D 1269 TTTATGCCAGAGAGATTTGACGAGTCTCTAAGGATTTTGTGTTAATAATTTTGAATAT 1328
QY 1286 GTACCATTTGGTCAGGAAAAAAGAAATATGTCCTGGCATTTACTTACGCTATTACCAATTTG 1345
D 1329 CTTCATTTGGTGGCGGAAGGAGGATTGTCCTGGGATTTGCTTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGTCATTATAATCTATTATATCATTTTAAATTGGGAACCTGCGCGATGGAATTACA 1405
D 1389 TATTGGCATTTGGCTCAATTTACTATATCATCTTCGATTGGAAACTCCCTGCTGGAATCGAA 1448
QY 1406 CCTCAAAACACTTGTATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAATAGATCTT 1465
D 1449 CCAAGCGACTTGGACTTGGACTGAGTTGGTGGAGTAACTGCGCTAGAAAAGTGCACCTT 1508
QY 1466 AAGTTGATTCCTATTCCATCAAA 1489
D 1509 TACTTGGTTGGGACTCTTATCAA 1532
```

RESULT 11

```
US-10-943-507-208
; Sequence 208, Application US/10943507
; Publication No. US20050132444A1
; GENERAL INFORMATION:
; APPLICANT: U. S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943.507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 208
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-943-507-208
```

```
Query Match      23.1%; Score 400.4; DB 10; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTATCATAGTAATCTTCCATTTACTTACTCTCT 172
D 135 CCACAGAGTCCATGGAACACTACCAATACTAGGAAGTATGCTTCATATGCT-----TGCT 188
QY 173 GATCTAGGCCATGAAGCTTTTTHAGAGCTTGGCTCAAAATTATGGACCTGTTATGAGTCTT 232
D 189 GGACTACCAACACCATGCTCTTAGAGATTTAGCAAAAAAATATGGAACCACTTATGCACCTT 248
```

```
QY 233 CAAATTGGCCAGTTTCAGCTGTGTGTCATTTCTTCTAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTTCGGGTTGGTTTACTTCTCTGTATACGCGCAAAAGATTA 308
QY 293 AAAACTCAGCTGATGCTTCGCCCAACGCCCTATCGTCTTGAGCCACAGATTGGTTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTCTAGCGCTTAGCTTTTGGCCCGGAGATTGCTGT 368
QY 353 TATAATCGGAAGATCTGTGTTGTTTCTATATGGAGATCACTGGAGGCGAGATGAAGAAA 412
Db 369 TACATAGTCTGATCTAGCTTTTGGCCCTATGCGCACTATTTGGAGACAATGCGTAAA 428
QY 413 ATTGGAATCTGAAATTTCTGATGCGCAAAAAAGTTCAATCTCTCAGGTTAATCCGAGAG 472
Db 429 ATATGTGTCTTGAAGTCTCAGTGCACAAGATGTTTCGGACATTTAGCTCTATTAGCGG 488
QY 473 GAAGAAATCGAGATGCCATCATTCTCCGTTTGGAAAGCCGGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCTGCTCTCAATTAATTTATCCGGTCACTCTTCTGGTGAACCTATTATGTT 548
QY 533 ACAAGATCAATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 AGGAAAGATCTTTTGTTCACAGCTCATGACATGATAGATCAGCGTTTGGCAAGTG 608
QY 590 TGTAAAGCAAAAAGATTGCTGAGTGTGCCGATGCACTCAATGAGGCGAGCGAGT 649
Db 609 TTTCAAGAGACAAGCAAAATTTATACAACTAAATTAAGAAAGTGATACTCTTAGCAGGAG 668
QY 650 TTTGGCACCAGACGCTTTTCGGAGCTGGAATTTACTTCACTATATCAATTTGAGCTGAG 709
Db 669 TTTGATGTGGCTGACATATTTCCCTTCACTGAAGTTTCTTCACTGTGCTCAGTGGAAATGA 728
QY 710 TCAAAACCCAGCGTTTGCATCAGGAGATTGAGATATATCTTGAAGAGATCTCTTAATGAA 769
Db 729 GGTAAATATGATGACACACCATAGGATAGTGCATTTGTTGAGATGTCATCAATGAG 788
QY 770 CAC-----AAAGCCAAATAAGCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAATGAGCGTTAGGAGTGAAGATTTAAT 848
QY 809 GATGTTCTATGAATCTTCAAAAAATGAAAGTTTCCAGTGCAGTGACAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGAGGCGCTTCAATTTCTTCACTCAACCAACGAC 908
QY 869 ATCAAGACATCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTTTCGAAAGCTACA 928
Db 909 ATCAAGACCAATATTTTTCACATGTTTGTGCGGAGCAGAGACTTCATCGTCAACAAT 968
QY 929 GAATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCACAAGAAA 988
Db 969 GTGTGGCTATGTGAGAAATGTTGAAAAATCCAGCGTATTTCGGAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTAGAGAGAGCAATTTAGAGAAAAAGAACTTTTCGATGAAAAATGATGTGGAGGAGCTAA 1088
QY 1049 TTTCTTCAAGTTAGTGTAAAGAACTCTAAGATTAACATCTCCCTCCGTTT---GCTCTGAT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGTCTC 1148
QY 1106 CCGAGGAGTGTAGAGAAACAAACAGAAATGTGATATGAATTTATCCGAACACTCGA 1165
Db 1149 CCAAGAGAAATGTAGGAAGAGACAAATATAAACGGCTACACTATTCTCTGTAAGACCAA 1208
QY 1166 ATTGTTGTGAATGCTTTGGCGATAGAGAGATTCCTAATACATTGTCGGAACCTGGAAG 1225
Db 1209 GTCATGGTTAAATGTTTGGGCTTTTGGGAAGAGATCCAAAATATTGGAATGACGCAAACT 1268
QY 1226 TTTAAACCCAGAAAGTTTAAAGATTTGCAATTTGATATTAAGGGAGCAGACATTTCACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGAATTTTGTGGTAATTAATTTGAAAT 1328
QY 1286 GTACCAATTTGGTGCAGGAAAAAGAAATATGTCTCTGGCAATTTACTTTCAGCTATTACCAATTTG 1345
```

```
Db 1329 CTTCCATTTGGTGGCGAAGGAGATTGTCTCTGGGATTTGTTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGCTAATTATAAATCTATATATCAATTTTAATTGGAACTGGCCCAAGTAATACA 1405
Db 1389 TATTTGCCATTGGCTCAATTTACTATATCACTTCGATTGGAATCCCTGCTGGAATCGAA 1448
QY 1406 CCTCAAAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGGACTTGAGCTTGACTGAGTTGGTGGAGTAACCTGCCCTAGAAAAAGTGACCTT 1508
QY 1466 AAGTTGATCTCTATTCCATATCAAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532

RESULT 12
US-10-934-944-209
; Sequence 209, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 209
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; US-10-934-944-209

Query Match 23.1%; Score 400.4; DB 11; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTTATCATAGTAGTAATCTTCTCTCAATTTACTCACTTCT 172
Db 135 CCACAGGTCCTATGGAATACCAATACTAGGAAGTATGCTTCATATGCT-----TGGT 188
QY 173 GATCTAGGCCATGACGTTTTAGAGCCTTTAGAGCCTCAAAATTTATGGACCTGTTATGAGTCTT 232
Db 189 GGACTACCAACCATGTCTTAGAGATTTAGCAAAAAAATATGGACCACTTATGACACCTT 248
QY 233 CAAATTGGCCAAAGTTTTCAGCTGTTTCTCATTTTCTTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTGCGGTTGGTGTACTTCTCTGATACGCGCAAAAGAGATTA 308
QY 293 AAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGAGCGCACAGATTGTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTCTAGGCTTAGCCCTTTTGGCCCCCGGAGATTGCTGT 368
```


Db 429 ATATGTCTCTTGAAGTGTCTCAGTGCACAAGATGTTCCGACATTTAGCTCTATTAGGCGG 488
QY 473 GAAGAAATGAGGATGCGATCATCATCTCCGTTTGAAGCCGATCTCCGGTCAATATT 532
Db 489 AATGAAGTCTTCGTCTCATTAATTTATCCGGTCACTTCTCGTGAACCTATAATGTT 548
QY 533 ACAAGATCATTTATGGCATTAATAATTTCCGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 ACGAAAGGATCTTTTGTGTTCAAGCTCATGATGATCAGCGTTTGGCGAAGTG 608
QY 590 TGTAAACAAAGAAAGATGCTGAGTGTTCGCGATGCGAGTCAATGAGGCGAGCGAGT 649
Db 609 TCAAGAGCAAGCAAAATTTATACAACTAATTTAAAGAAAGTGATCTCTTAGCAGGAGG 668
QY 650 TTTGGCACCAGCAGCGCTTTCCGAGCTGGAAATTTACTTCACTATATCATTTGAGCTGAG 709
Db 669 TTTGATGTGCTGACATATTCCTTCACTGAAATTTCTTCATGTGCTCAGTGAATGAAG 728
QY 710 TCAAAACCCAGGCGTTTGCATCAGGAGATGAGCATATATCTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCAATAAGGTAGATGCCAATGTTGAGAAATGTCATCAATGAG 788
QY 770 CAC-----AAAGCCATAAGCCTTTTGAAGCCGATTAATG 808
Db 789 CACAAGAAATCTTGCAATTTGGGAAAGCTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
QY 809 GATGTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAAGC 868
Db 849 GATGTTCTTGAAGCTTATGAATGATGAGGCGTTTCAATTTCTTATCCACGACAAAC 908
QY 869 ATCAAGCATCCGTTTTCGAAATGTTTACTGCGGAGGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAGCCATAATTTTGCATGTTTGTGCGGAGCAGAGACTTTCATCGTCAACAAT 968
QY 929 GAATGGTTAATGCGCAGAGCTGATGAATAATCCAACTGAACTAAGAAAGCAAGAAGAA 988
Db 969 GTGTGGCTTGTGTAATAATGTTGAAATAATCCAGCGGTATTCGGAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGCGTGAATGCGAAAGTGTGATGAATCAAGATTTTCATGATTGAAA 1048
Db 1029 GTAAGGAAGCATTTAGAGAAAGAAAGAACTTTCGATGAATAATGATGAGGAGCTAAAC 1088
QY 1049 TTCTTCAAGTTAGTGTAAAGAACTCTAAGATTACATCTCCCGGTT---GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTATAAAGAAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGGTAGAGAAACAACAGAAATGATGATGAATAATTCATCCGAACTCGA 1165
Db 1149 CCAAGAGAAATGAGGGAAGAGACAAATATAAAGCGGTACACTATTCTGTAAAGACCAA 1208
QY 1166 ATTTGTTGAATGCTTTGGGCGATAGGAAGAGATCCTAATCTTGGTCGGAACCTGGAAG 1225
Db 1209 GTCATGTTAATGTTTGGGCTTTGGAGAGATCCAAATAATTTGGAATGAGCGAGAACT 1268
QY 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGAGCAGCATTTGAAC 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTTGTGGTAATAATTTGAATAT 1328
QY 1286 GTACCAATTTGGTCAGGAAAAAGAAATATGTCCTGGCAATTAATTCAGCTAATACCAAT 1345
Db 1329 CTTCCATTTTGGTGGCGGAGGAGGATTTGTCCTGGGATTTGCTTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGCTATTATAATCTATTATATCAATTTTAAATTTGGAACTGCGCGATGGAATTACA 1405
Db 1389 TATTTGCCATTTGCTCAATTTACTATATCACTTCGATTTGGAATCTCCCTGCTGGAATCGAA 1448
QY 1406 CCTCAAAACATTTGATAGTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAATAGATCTT 1465
Db 1449 CCAAGGCACTTGAGCTTGACTGAGTTGGTTGGAGTAATGCGCGCTAGAAAAAGTGACCTT 1508
QY 1466 AGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTGGCGACTCCTTATCAA 1532

RESULT 14
US-10-934-944-231
; Sequence 231, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/341008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 231
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-934-944-231

Query Match 23.1%; Score 400.4; DB 11; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
QY 113 CCTCAGGCGCCATGGAAGTTTCCTATCATAGTAACTCTCCCTCAITTTATTACTCACTCTT 172
Db 135 CCACCAGGTCCATGGAAACTACCAATACCTAGGAAGTATGCTTCATATGTT-----TGGT 188
QY 173 GATCTAGGCGCATGAACGTTTTAGAGCCTTTGGCTCAAATTTATGGACCTCTTATAGCTCTT 232
Db 189 GGACTACCAACCATGTCTTAGAGATTAGCCAAATAATATGGACCCTTATGCACCTT 248
QY 233 CAAATGGCCAAAGTTTTCAGCTGTTGTCAATTTCTCAGCTGAAGAGCGCCAAAGAGTTATG 292
Db 249 CAAATAGGTGAAGTTTCTCGGTTTGTGGTTACTTCTCTGATACGCAAAAGAAATATTA 308
QY 293 AAAAATCAGGCTGATGCGCTTCCGCCCAACGCGCTATCGTCTTCGACGACAGATTTGTTT 352
Db 309 AAAATCTATGACATCGCTTTTGGCTTAGGCTTAGGCTTTTGGCCCCGAGATTTGCTGT 368
QY 353 TATAATCGAAAGATGTCTTGTTCCTCATATGGAGATCACTGGAGGCGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATAGCTAGCCTTTTGGCCCTATGGGACTATTGGAGACAAATGCGTAA 428
QY 413 ATTTGGATACATTGAATTTCTGAGTCCCAAAAAAGTTTCAATTCCTCAGGTTAATCCGAGAG 472
Db 429 ATATGTGTCTTGAAGTGTCTCAGTCCCAAGAAATGTTTCGACATTTAGCTCTATTAGCGG 488
QY 473 GAAGAAATGAGGATGCCATCACATTCCTCCGTTCCAAAGCGGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGTTCTCATTAATTTTATCCGGTCACTTCTTGGTGAACCTTAAATGTT 548
QY 533 ACAAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589

```
Db 549 ACGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGCAAGTG 608
Qy 590 TGTAGCAAAAGAAAGATTGCTGAGTGTGCGGATGAGTCAATGAGCGAGCGAGT 649
Db 609 TTCAAGAGCAAGCAAAATTTATACAACTAATTAAGAAGTGATCTTTAGCAGGAGG 668
Qy 650 TTGCGACCGCAGACGCTTTTCCGACGCTGGAAATTTACTCATATATCATTTGGAGCTGAG 709
Db 669 TTTGATGCTGACATATTCCTTCACTGAAGTTTCTTCATGTCTCAGTGAATGAAG 728
Qy 710 TCAAAACCCAGCGCTTTCATCAGAGATTTAGAGATATCTTTGAAGAGATCTTAAATGA 769
Db 729 GGTAAAGATTATGAATGCACCATAGGTAGTGCATTTGAGAAATGTTCATCAATGAG 788
Qy 770 CAC-----AAAGCCATTAAGCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CAAAGAAAAATCTTGGAAATTAATGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGGAACGTTTCCAGTGCCAGTGACAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGAGGAGCTTCAATTTCTTATCACCACGACAA 908
Qy 869 ATCAAGACATCCGTTTTCGAAATGTTTACTGCGGAGGAGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAGCCATTAATTTTGAATGTTTGTGTCGCGGACAGAGACTTTCATCGTCAACAAT 968
Qy 929 GAATGGTAAATGCGAGAGCTGATGAAATCAACTGAACAAAGAAAGACAAAGAA 988
Db 969 GTGTGGCTATGTTGAATGTTGAATATCCAGCGGATTTCCGAAAGCTCAACAGAA 1028
Qy 989 GTTAGCAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTAGAGAGACATTTAGAGGAAAGAAACTTTTCGATGAATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TTTCTCAAGTTAGTGTGTAAGAAACTCTAAGATTTACATCTCCCGTT---GTCTGATT 1105
Db 1089 TACCTAAAGTTAGTAAATTAAGAAACTCTAAGACTTTATCCACCGGTTCCACTTTTGCT 1148
Qy 1106 CCGAGGAGTGTAGAGAAACAAACGAAATTTGATGATATGAATTTATCCGAAACACTCGA 1165
Db 1149 CCAAGAGATGTAGGAGAGACAAATATAACGCTACACTATTTCTGTGAAGACCAA 1208
Qy 1166 ATTGTTGTAATGCTTGGCGGATAGAGAGATCTTAATATCTTGTGCGAACTCGAAAG 1225
Db 1209 GTCATGTTTAATGTTTGGGCTTTGGAGAGATCCAAATATTTGGAATGACGAGAACT 1268
Qy 1226 TTTAACCCAGAGGTTTAAAGATTTGCAATTTGATTAAGAGGAGACATTTGCACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTCTTAAGGATTTTGTGGTAAATTTTGAATAT 1328
Qy 1286 GTACATTTGGTGCAGGAAAGAAATATGTCCTGCGATTACTTTCAGCTATTACCAATTG 1345
Db 1329 CTTCAATTTGGTGGGAGAGGATTTGCTCTGGGATTTCTGCTTACCTAATGCT 1388
Qy 1346 GAGTATGCTAATTAATCTATTATATCTTTTAAATTTGGAATCTGGCGATGGAATTA 1405
Db 1389 TATTTGCCATTTGGCTCAATTAATCTATATATCTTTCGATTTGGAATCTCCCTGCTG 1448
Qy 1406 CCTCAACACTTGTATGATGATGAGCTATTGGCGGTCTCTCAGGAAAGAAATAGATCT 1465
Db 1449 CCAAGCGACTTGGACTTGAATGAGTTGGTGGTAACTGCGGCTAGAGAAAGTGAACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532
```

RESULT 15
US-11-116-881A-218
; Sequence 218, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:

```
; APPLICANT: Xu, Dongmei  
; APPLICANT: Nielsen, Mark T.  
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof  
; FILE REFERENCE: 07678/141014  
; CURRENT APPLICATION NUMBER: US/11/116,881A  
; CURRENT FILING DATE: 2005-04-27  
; PRIOR APPLICATION NUMBER: 60/665,451  
; PRIOR FILING DATE: 2005-03-24  
; PRIOR APPLICATION NUMBER: 60/665,097  
; PRIOR FILING DATE: 2005-03-24  
; PRIOR APPLICATION NUMBER: 60/646,764  
; PRIOR FILING DATE: 2005-01-25  
; PRIOR APPLICATION NUMBER: 60/607,357  
; PRIOR FILING DATE: 2004-09-03  
; PRIOR APPLICATION NUMBER: 60/566,235  
; PRIOR FILING DATE: 2004-04-29  
; PRIOR APPLICATION NUMBER: 10/934,944  
; PRIOR FILING DATE: 2004-09-03  
; PRIOR APPLICATION NUMBER: 10/943,507  
; PRIOR FILING DATE: 2004-09-17  
; PRIOR APPLICATION NUMBER: 60/503,989  
; PRIOR FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: 60/485,368  
; PRIOR FILING DATE: 2003-07-08  
; PRIOR APPLICATION NUMBER: 60/418,933  
; PRIOR FILING DATE: 2002-10-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2300  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 218  
; LENGTH: 1673  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-11-116-881A-218
```

```
Query Match 23.1%; Score 400.4; DB 16; Length 1673;  
Best Local Similarity 57.7%; Pred. No. 1.2e-88;  
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;  
  
Qy 113 CCTCAGGCGCATGGAAGTTTCTATCATAGGTATCTTCTCATTTTACTCTCACTTCT 172  
Db 135 CCACAGGTCCATGGAACTACCAATCTAGGAAGTATGCTTCATATGTT-----TGGT 188  
  
Qy 173 GATTCAGCCATGAACGTTTGTAGAGCCTTGGCTCAAAATTTATGAGCTTGTATGAGTCTT 232  
Db 189 GGACTACCAACCATGCTCTTAGAGATTTAGCCAAAAAATATGGACCACCTTATGCACCTT 248  
  
Qy 233 CAAATTGGCCAAAGTTTTCAGCTGTTGTCTTCTCAGCTGAAGCAGCCAAAGAGGTATG 292  
Db 249 CAATTAGGTGAAGTTTCTGCGGTTGTGTTACTTCTCTGATACGCGCAAGAGATTTA 308  
  
Qy 293 AAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGAGCGCACAGATTGTGTTT 352  
Db 309 AAACTCATGACATCGCTTTTGGCTTAGGCTAGGCCCTTTTGGCCCCCGAGATTGTCTGT 368  
  
Qy 353 TATTAATCGAAAGATGCTTGTGTTGCTTCATATGAGATCACTGAGCGCAGATGAAGAA 412  
Db 369 TACAATAGTCTGATCTAGACCTTTTGGCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428  
  
Qy 413 ATTTGGATCTTGAATTTCTGAGTCCCAAAAAAGTTTCAATCTCTCAGGTTTAAATCCGAGAG 472  
Db 429 ATATGTGTTGGAAAGTCTCAGTGCCCAAGAAATGTTTCGGACATTTAGCTCTATTAGGCGG 488  
  
Qy 473 GAAGAAATGAGGATGCCATCACATTCCTCGTTCGAAAGCGGATCTCCGGTCAATATT 532  
Db 489 AATGAAGTCTTCTGCTCATTAAATTTATCCGGTCATCTTCTGTGTAACCTATTAAATGTT 548  
  
Qy 533 ACAAGATCATTTATGGCATTTAATTTTCGATCATGATAAGACATCCGTTTCTGTA---T 589  
Db 549 ACGGAAAGGATCTTTTGTTCACAAAGCTCCATGATGTAGATCAGCGTTTGGGCAAGTG 608  
  
Qy 590 TGTAAAGCAAAAGAAAGATTTGCTGAGTGTGTCGAGTGCATGAGTCAATGAGGACGAGCT 649
```

Db 609 TTCAAAGACGACACAAATTTATACAACTAAATTAAGAAGTGATACCTCTTAGCAGGAGG 668
Qy 650 TTGGCACCAGCAGACGCTTTTCGACGTGGAATTAATTTCACTATATCATTTGGAGCTGAG 709
Db 669 TTTGATGTGGTGACATATTCCTTCACTGAAGTTTCTCATGTGCTCAGTGAATGAAG 728
Qy 710 TCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTTTGAAGAGATTCTTAAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCAATAAGGTAGATGCCATTGTTGAGAATGTCATCAATGAG 788
Qy 770 CAC-----AAAGCCATTAAGCCTTTTGAAGCGGATAACTTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGAAAACTTAATGGAGCGTTAGGAGGTGAAGATTTAATT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGAAACGTTCCAGTCCAGTGACGACAAACGAAGC 868
Db 849 GATGTTCTTAAGACTTATGAATGATGAGGCGTTCAATTTCTATACCCACGACAAAC 908
Qy 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAAGCTATAATTTTGACATGTTTGTGCGGGACAGAGACTTCATCGTCAACAATT 968
Qy 929 GAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACCTGAAGAAAGCAAGAAGAA 988
Db 969 GTGTGGGCTATGTTAGAAATGGTGAATAATCCAGCCGTATTCGCGAAAGCTCAAGCAGAA 1028
Qy 989 GTTAGACAGATTTGGTGAATGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTPAAGAGAGCAATTTAGAGGAAAGAACTTTTCGATGAAAAATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TTCTTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTTAAGACTTCATCCACCGTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGGTAGAGAAACAACAGAAATGATGGAATGAAATTCATCCGAACACTCGA 1165
Db 1149 CCAAGAGAATGTAGGGAAGAGACAAATATAAACGGCTACACTATTCTGTAAAGACCAA 1208
Qy 1166 ATGTTGTGAATGCTTTGGSCGATAGGAAGATCCTAATACCTTGGTCGGAACCTGGAAG 1225
Db 1209 GTCATGGTTAATGTTTGGCTTTGGGAAGAGATCCAAATATTTGGAATGACCGAGAACT 1268
Qy 1226 TTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAAGAGGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTGAGCAGTGTCTAAGGATTTTGTGTTAATAATTTTGAATAT 1328
Qy 1286 GTACCATTTGGTCAGGAAAAAGAAATATGCTCTGGCATTTACTCAGCTATTACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGCGGAAGGAGATTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACTGGCGATGGAATTACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTAATCTATATCACTCGAATTTGAAACTCCCTGCTGGAATCGAA 1448
Qy 1406 CTTCAAAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCGACTTGACTTGACTGAGTTGGTGGAGTAACTGCCGCTAGAAAAAGTACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCCTTATCAA 1532

Search completed: May 31, 2006, 10:17:07
Job time : 2317 secs

This Page Blank (uspto)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: May 31, 2006, 09:38:38 ; Search time 35 Seconds
(without alignments)
5831.494 Million cell updates/sec
Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gcataaaggaaatggagc.....tcttttcaatccgaaaaa 1733
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 246937 seqs, 5886990 residues
Total number of hits satisfying chosen parameters: 493674

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_New.*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB_seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB_seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB_seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB_seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	384.2	22.2	1751	6	US-10-953-349-22681 Sequence 22681, A
2	367.4	21.2	1504	6	US-10-953-349-24167 Sequence 24167, A
3	248.4	14.3	1737	6	US-10-953-349-14590 Sequence 14590, A
4	224.2	12.9	1915	6	US-10-953-349-13595 Sequence 13595, A
5	216.4	12.5	1617	6	US-10-953-349-2239 Sequence 2239, Ap
6	195.4	11.3	1637	6	US-10-953-349-6466 Sequence 6466, Ap
7	172.8	10.0	1958	6	US-10-953-349-17424 Sequence 17424, A
8	166	9.6	1586	6	US-10-953-349-4847 Sequence 4847, Ap
9	159	9.2	1659	6	US-10-953-349-13112 Sequence 13112, A
10	158.4	9.1	1658	6	US-10-953-349-17280 Sequence 17280, A
11	153.8	8.9	1722	6	US-10-953-349-10369 Sequence 10369, A
12	152.2	8.8	1660	6	US-10-953-349-21323 Sequence 21323, A
13	148.4	8.6	1632	6	US-10-953-349-32482 Sequence 32482, A
14	148.2	8.6	1766	6	US-10-953-349-32468 Sequence 32468, A
15	147.2	8.5	1727	6	US-10-953-349-12509 Sequence 12509, A
16	137	7.9	1639	6	US-10-953-349-32472 Sequence 32472, A
17	130.6	7.5	1711	6	US-10-953-349-33880 Sequence 33880, A
18	124.6	7.2	1720	6	US-10-953-349-32176 Sequence 32176, A
19	116	6.7	1636	6	US-10-953-349-21349 Sequence 21349, A
20	110	6.3	1852	6	US-10-953-349-12597 Sequence 12597, A
21	103.2	6.0	1667	6	US-10-953-349-2736 Sequence 2736, Ap
22	102.6	5.9	1866	6	US-10-953-349-13642 Sequence 13642, A
23	100.8	5.8	1890	6	US-10-953-349-12761 Sequence 12761, A
24	99.8	5.8	1165	6	US-10-953-349-22519 Sequence 22519, A
25	99.2	5.7	1756	6	US-10-953-349-13063 Sequence 13063, A

ALIGNMENTS

RESULT 1
US-10-953-349-22681
; Sequence 22681, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22681
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-22681

Query Match	22.2%	Score 384.2;	DB 6;	Length 1751;
Best Local Similarity	56.6%	Pred. No. 1.3e-92;		
Matches 813;	Conservative 0;	Mismatches 588;	Indels 36;	Gaps 4;
QY	104	CAGAAATCCACCTCCAGGGCCATGGAAGTTTCTCTATCATAGGTAATCTTCTCATTTTATTA	163	Sequence 32215, A
DB	133	CACAAACTACCCCGGACCAAGAAGCTACCTCTGATAGGGAACCTGCACTCACTAGCA	192	Sequence 9959, Ap
QY	164	CTCATTCTGATCTAGGCATGAACGTTTTAGAGCCTTGCGCTCAAAATTTATGGACCTGTT	223	Sequence 32409, A
DB	193	ATGGCTGGTTTCACTTCCACATCGTACTCTCCGAGACCTTGGCCCTTAAATACGGACCTCTC	252	Sequence 3128, A
QY	224	ATGAGCTTCAAAATGGCCAGTTTTCAGCTGTTGTTCATTTCTTACGTCGAAGCAGCCAAA	283	Sequence 31583, A
DB	253	ATGCAATCCAACTGGTGAATTTTCTCCGGTGTGTATCTCCCCCAACATGGCCAAAG	312	Sequence 34955, A
QY	284	GAGGTATGAAAACCTCAGGCTGATGCCCTTCGCCCAACGCCCTTCGTCTTTGGACGCAAG	343	Sequence 18767, A
DB	313	GAGATAATGAAACCCATGATCTTCTTTTGTGACAGAGCCACAAATTTCTCTCTCTCAA	372	Sequence 25760, A
QY	344	ATTGTTTTTAAATCGGAAAGATGCTCTTTTGTCTTCTATGGAGATCACTGGAGGAG	403	Sequence 32448, A
DB	373	ATTTTGAACCTATGGGCAAAATGATATGTTTTTGTCTTCCATATGGGGATTTATGGAGCAA	432	Sequence 4124, Ap
QY	404	ATGAAGAAAATTTGGATACCTTGAATTTCTGAGTGCACAAAAGTTTCAATCTCTCCAGTTA	463	Sequence 34350, A
DB	433	ATGAAGAAAATATGTTGTCGGAGCTTCTCGGAGCTTCTTAAGTGCACAGAGAGTTTCAGTCTTCTCAT	492	Sequence 26798, A
QY	464	ATCCGAGAGGAGAAATGGAGGATGCCATCACATTTCTCCGTTTCGAAAGCCGGATCTCCG	523	Sequence 32381, A

Db 493 ATTGAGAAGACGAGACATCCAAATTTATAGAATCAATTCGAATATCAAGAGGTTACCA 552
QY 524 GTCATATTTACAAAGATCATTTATGGCATTAATTTTCGATCATGATAAG--AACATCC 580
Db 553 ATCAATCTTACTAGTAAATTTACTCCTTGGTAGTCTTCTGTTCCAGGGTAGCATTT 612
QY 581 GTTGCTAATTTGTAAGCAAAAGAAAGATGCTGAGTGTGCGGATGCGAGTCAATGAGGCA 640
Db 613 GCGACAAATCTAAGGACCAAGAGATTTTGTGTGCTTGAAAAAATGATATTAGCG 672
QY 641 GCGAGAGTTTGGCACCAGCAGACGCTTTTCGAGCTGCGAAATTTACTTCATATCATTT 700
Db 673 GGGGAGGATTTGAACCTGACGATTTGTTTCCCTCGATGAA---ACTTCATCTCATAAAT 729
QY 701 GGAGCTGAGTCAAAACCCAGGCGTTTGATCAGGAGATTTGAGATATATCTTGAAGAGATT 760
Db 730 GGAAGGAAGGCCAAATTTGGAGAGATGATGAGCAGGTTGACAGATCGCAGACACATTT 789
QY 761 CTTAATGAACACAAAGCCCAATAAGCCTT-----TTGAA 793
Db 790 CTCAGAGGACATCAAGAAAGCGCCAAAGGCGCACTACGAGAAGGCAAGTTGACCTAGAG 849
QY 794 GCGGATACTTAATGGATGTTCTATTGAACTTTCAAAAAATGGAAGCGTTCCAGTGCCA 853
Db 850 GAGGAGGATCTTGTGTGATGTTCTTTTGAGAAATCCAAACAAAGTGACAACCTCGAGATCAA 909
QY 854 GTGCAACAGGAAGCATCAAGCATCCGTTTGGCAAAATGTTTACTGCGGAGCGCAACA 913
Db 910 ATTAGTACACANACATCAAGCTGTTATATTGGACGTTATTACTGCTGGAAACGATACT 969
QY 914 ACTTCGAAGCTPACAGAATGGGTAAATGCGAGAGCTGATGAAAAATCCAACTGAATAGA 973
Db 970 TCAGCATCAACACTTGAATGGCTATGGCAGAAATGATGAGAAACCCAAAGAGTGAGGAAG 1029
QY 974 AAAGCACAAGAAAGATTAGACAAGTATTGTTGAAATGGGAAAGTTGATGTAATCAAGA 1033
Db 1030 AAAGCACAGGCTGAAGTAAGACAAGCATTTAGGGAATTTGAAATTAATCCATGAACACAGAT 1089
QY 1034 TTTTCATGATTTGAAATCTTCAAGTTAGTGTGTTAAAGAACTCTAAGATTACA---TGCT 1090
Db 1090 GTAGGGAACCTTACATCTTAAGTTGGTGATCAAGAGACATTTAAGGTTTACACGCTCT 1149
QY 1091 CCGGTTGTCTTGATTTCCGAGGAGGTGTAGAGAAACAACAAGTAATTTGATGGATGAAATTT 1150
Db 1150 TCTCCTTTATTGGTCCCTAGAGAAATGCTCTGAACCTAACCATCATTTGATGGATGAAATA 1209
QY 1151 CATCCGAACACTCGAATTTGTTGTAATGCTTGGCGGATAGGAGAGATCCCTAATACTGG 1210
Db 1210 CCTGTCAAGACTTAAAGTCATGATNAACCGTGTGGGCAATTTGGAAGAGATCCCCCAATTTGG 1269
QY 1211 TCGGAACCTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGG 1270
Db 1270 ACTGATCTGAGAGGTTTGTCCGAGAGATTCGATGGTAGTCTTATCGATTTCAAGGG 1329
QY 1271 ACGACATTTGAACTGGTACCTACCTATTTGGTGAGAAAGAAATATGCTTGGCAATTACTTCA 1330
Db 1330 AATAACTTTGAGTATCTTCTCTTTGGGGCAGGAAGAGAAATGTGCCAGGAATGACATTT 1389
QY 1331 GCTATTACCAATTTGGAGTATGTCATTATAATCTATTATATCAATTTAATTTGGGAACCTG 1390
Db 1390 GGTTTAGCCCAACATTTATGCTTCCACTCGCTCTATTACTCTATCACTCAATTTGGGAACCTT 1449
QY 1391 GCGGATGGAATTTACACCTCAAAACACTTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGG 1450
Db 1450 CCAATGAGATGAACCTGAGGATATGGATATGTCTGAAAACTTTGGATTGACAGTTACA 1509
QY 1451 AAAAAAATAGATCTTAAGTTGATTCTTATTTCCATATCAAGTTAGCTTTAGGCTCAAT 1507
Db 1510 AGGAAAAGCGAGTTGTGTTTGATTCCCATTTGTAATGATCTTTGACACTGCAACGATT 1566

US-10-953-349-24167
; Sequence 24167, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24167
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-24167

Query Match 21.2%; Score 367.4; DB 6; Length 1504;
Best Local Similarity 56.8%; Pred. No. 3.5e-88;
Matches 752; Conservative 0; Mismatches 546; Indels 27; Gaps 3;

QY 191 TTTAGAGCCTTGGCTCAAATTTATGGACCTGTTATGAGTCTTCAAATTTGGCCAACTTTCA 250
Db 9 TTAAGAGACCTTGGCATATAAATATGGACCGTGTGATGATCTTCAACTTTGGACAGTTTCA 68
QY 251 GCTGTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATGAAAACTCAGGCTGATGCC 310
Db 69 ACTATTGTGATTTTCATCCCTGAGTGTGCTAGGGAAGTGATGAAACCCATGACATTAAC 128
QY 311 TTCCGCCAACCCCTATGCTCTTGGACGCACAGATTTGTTTATTAATAATGCGGAAGATGTC 370
Db 129 TTTGCCACAAGGCCCTTAAAGTTCTAGCTATTGAAATAATGTCTTACATTTCCACAAGTATA 188
QY 371 TTGTTTGTCTCATATGGAGATCACTGGAGGAGATGAAGAAAAATTTGGATCTTGAATTT 430
Db 189 GCTTTTGTGCTATGGAATTTATGGAGGAGCTAAGAAAAATATGACATTTGGAGCTT 248
QY 431 CTGAGTGCCAAAAAGTTCAATCTCTCAGGTTTAACTCCAGAGAGGAAGAAATGGAGGATGCC 490
Db 249 TTAAGCCTTAAACCGCTCAACTCTTCCAGCCAATTAGAGAAGATGAGCTCTTCAATCTT 308
QY 491 ATCACTATCTCCGTTTCGAAAGCCGGATCTCCGGTCAATATTTACAAAGATCATTTATGCG 550
Db 309 GTCAATGGATGATTCAAGAAAGGATCCCCCATCAACCTCCTGAGGAGTACTTACA 368
QY 551 ATTATAATTTCCGATCATGATAAGAACATCCGTTGG---TAATTTGTAAGCAAAAGAGA 607
Db 369 TCAATTTACAAATTTGCTTCAAGGGCTGCCCTTTGGCAAGNATTTGCAAGACCAAGAAAA 428
QY 608 TTGCTGAGTGTTCGGATGCGAGTCAATGAGGAGCGACGAGTTTGGCACCCAGCGCT 667
Db 429 TTTATATCAGTGTGTTAAAAAACAATCAAAACTTGTCTGCAGGTTTGGCAATTGAAGATTG 488
QY 668 TTTCCGAGCTGGAATTTACTTCACTATATCATTTGAGCTGAGTCAAAACCCAGGGCTTTG 727
Db 489 TTTCTCTTCTGTTACTTTGGCTTCAACATGTCTACTGCTTGGAGGCTTAAGCTTGAAGGTTG 548
QY 728 CATCAGGAGATTGAGATATATCTTGAAGAGATTTTAAATGAACAACAAGCCCAATAGCCT 787
Db 549 CATCAACAGGCTGATCAGATAATGAAAAACATCATCAATGAGCATAAAGAGGCAAAATCG 608
QY 788 TTTGAAGCGGATAACTTAA-----TGATGTTCTTATGATCTT 826
Db 609 AAAGCAAGGATGACCAAAAGTGAAGCAGAAGAAGATCTTTGTGGATGTTCTCATACAATAT 668
QY 827 CAATAAATGGAACGTTCCAGTGCAGTGACAAAACGAAAGCATCAAAGCATCCCGTTTGG 886
Db 669 GAGGATGGAGCAGAGAGGATTTTCTTAACTAGAAACAGATCAAGGCCAATATTCTG 728
QY 887 CAAATGTTTACTCCGGGAGCGAAAACAACTTCGAAAGCTACAGAATGGGTAATGGCAGAG 946
Db 729 GACATTTTGTGCTGTGGAGGTGAGACAAACAGCAACCAACCATAGATTGGGCAATGGCTGAA 788

QY 947 CTGATGAAATCCAACTGAACCTAAGAAAGCAAGAAAGCTTAGACAAGTATTGGT 1006
 DB |||||
 DB 789 ATGGTAAAGATCCACAGTAAAGAAAGCAACATCTGAGTGAGAGGTATGCAAT 848
 QY 1007 GAAATCGGAAAGTGTGATGAATCAAGATTCATGATTTGAAATCTTCAAGTTAGTGGT 1066
 DB |||||
 DB 849 ATGAAAGCAAGGTTGATGAAATTCATCAATGAACCTTCAATATTTGAAACTTAATTTGTG 908
 QY 1067 AAGAAACTCTAGATTTACATCTCCGG---TTGTCTTGATTCGGAGGAGTGTAGAA 1123
 DB |||||
 DB 909 AAGAGACCTTGAGGTGACCCCTCCAGCTCCTCTTTTGTGCTCAAGAGAATGTGTCAA 968
 QY 1124 ACAACACGAATTCATGGATATGAATTCATCCGAACACTCGAATTTGTGTGAATGCTTGG 1183
 DB |||||
 DB 969 ACATGTGAGATACATGTTATCATACATACCAGCCAAACCAAGTCAATGCTGCTGG 1028
 QY 1184 GCGATAGGAGAGATCCTTAATCTTGGTGGAACTCTGGAAAGTTTAAACCCAGAAAGTTT 1243
 DB |||||
 DB 1029 GCAATTTGGAAGAGATCCAACTATTGGAATTCGAACTGAATCAGAGAGGTTTATCTGAGAGGTTT 1088
 QY 1244 AAGATTTGCAATTCATTAAGGAGGACGACATTTGAATGCTGTACCATTTGGTGCAGGA 1303
 DB |||||
 DB 1089 ATTGATGACCACTATTGACTACAAAGGAGTAATTTTGAATTCATCTCTTTGGTGTGCTGA 1148
 QY 1304 AAAAGAAATATGCTCGCATTAATTCAGCTATTACCAATTTTGGAGTATGTCATTAATAAT 1363
 DB |||||
 DB 1149 AGAAGATATGCGAGGAGACATTTGCTTTGAGAGCTCGGAACTAGCCCTTGCATG 1208
 QY 1364 CTATTATATCATTTTAATTTGGGAATGCGCGGATGGAATTAACCTCAACACATTTGATATG 1423
 DB |||||
 DB 1209 TTGTTGTATCACTTTGATTTGGAAGCTTCCAAAGTGAATGAGAAGTGGAGAACTGGACATG 1268
 QY 1424 ACTGAAGCTATTGGCGTCTCTCAGGAAATAATAGATCTTAAGTTGATCTCTATTCCA 1483
 DB |||||
 DB 1269 AGTGAAGATTTGGAGTCAACTATAAGAAAGATAATCTATCTTTGTTGCTCTTTTCCA 1328
 QY 1484 TATCA 1488
 DB |||||
 DB 1329 TATCA 1333

RESULT 3
 US-10-953-349-14590
 ; Sequence 14590, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE OF INVENTION: ENCODED THERBY
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 14590
 ; LENGTH: 1737
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-953-349-14590

Query Match 14.3%; Score 248.4; DB 6; Length 1737;
 Best Local Similarity 51.4%; Pred. No. 1.1e-56;
 Matches 722; Conservative 0; Mismatches 646; Indels 36; Gaps 5;

QY 48 TAAATAGTTTCTGCTGTTGTTAACTTTAGTAGTAGTCAATGAGGTTGTGGAAGAAACAGA 107
 DB |||||
 DB 171 TAACTCTCTTTGTTGTTGTTCTTCTTCAATACCGTAAAAACCTTCAAGAAATCCAC 230
 QY 108 ATCCACCTCCAGGCCATGGAAGTTTCCCTATCATAGGTAATCTCTCTCATTTATTACTCA 167
 DB |||||
 DB 231 CACTTCCACCAAGGCTCTAGAGCCCTTCCCAATAATAGGGAATCTTCAATCAATAGATGTC 290

QY 168 CTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTGTTTATGA 227
 DB |||||
 DB 291 CTTCTCTT-----CATGAGCAGCTATGCACTCTCTCAAGAAATATGCTCTTTATTTT 344
 QY 228 GTCTTCAAAATCGCCAAAGTTTCAAGCTGTTGTCTATTTCTTTCAGCTGAAGCAGCCAAAGAGG 287
 DB |||||
 DB 345 CCCTTCAATTTGGATTAAGGCCAGCCATAGTTGTTGTTCTCTCCAAATTTGGCCAAAGAGG 404
 QY 288 TTATGAARAACTCAGGCTGATGCTTCCGCCCAAGCCCTATCGTCTTTGGAGCGCACAGTTG 347
 DB |||||
 DB 405 TAATGAAGAACAATGACCTTGAGTGTGTGGAGCAGCTAAATTTACTAGGCCAACAGAAAC 464
 QY 348 TGTTTTATAATCGGAAGAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGCGCAGATGA 407
 DB |||||
 DB 465 TAACTATAATGGTTAGACATGGGATTTTCTCTCATACGATAGTTATTTGGAGAGAATCA 524
 QY 408 AGAAATTTGGATATCTTGAATTTCTGAGTGCCAAAGAGTTTCAATCTCAGGTTTAAATCC 467
 DB |||||
 DB 525 GAAAAATTTGTGTTGCTCATGCTTAGCTCCAAACGCTGTCCAAAGCTTTTACCTCAATAA 584
 QY 468 GAGAGGAAGAAATGGAGGATGCCATCACATCTCTCGTTTCGAAAGCCGATCTCC----- 522
 DB |||||
 DB 585 GACATTTGAGGTCAAGCAGATGATAAAAAAATATCTAGGCATGCTCATCTTCAAAAG 644
 QY 523 -GGTCAATATTACAAAGATCATTTATGGCATTATATATTTTCGATCATGATAAGAACATCCG 581
 DB |||||
 DB 645 TTACAAATTTAATGAAGTCTTATATCCCTTATTAGTACTATTGTATGTAGAAATTTGTTT 704
 QY 582 TTGGTAAATTTGAAGCAAAAGA-----AAGATTGCTGAGTGTGTCGCGATG 626
 DB |||||
 DB 705 TGGGAGAGAGGTATGAAGAGGAGGATCTGAAGGTAGCAGATTTCCATAAAGTTGTTCAATG 764
 QY 627 CAGTCAATGAGGACGCGAGGTTTGGCACCAGCAGCGCTTTCCGACGCTGCGAAATTAAC 686
 DB |||||
 DB 765 AGTGTGAAGCTATGTTGGTAACTTCTTTGTTTTCAGATTATATTTCTTTTCAATGGGTTGGA 824
 QY 687 TTCCTATATATCAATGGAGCTGAGTCAAAACCCAGGCGTTTGTGATCAGAGGATTTGACGATA 746
 DB |||||
 DB 825 TTGATAAACTCAGGGGACTGGATGACGCTTTGMAAGCAATTTCAAGGAGATGGAATAAGT 884
 QY 747 TACTTGAAGAGATTTCTTAATGAACACA-----AAGCCAAATGAAGCTTTTGAAGCCGATA 800
 DB |||||
 DB 885 TCTACCAAGAAGCCATTTGATGAACACATGAATTTCTAAGAAATAAATACTCCAGAGGAAGAGG 944
 QY 801 ACTTAATGATGTTCTTATTGATCTTCAAAAAATGGAACCTTCCAGTCCGACGATGACAA 860
 DB |||||
 DB 945 ATTTAGTTGATGCTTTACTTCAACTGAAAGAGAACAAACGTTTCCCATAGATCTCACCA 1004
 QY 861 ACGAAAGCATCAAAAGCATCCGTTTTCGAAATGTTTACTCCGGGAGCGAAACAACTTCGA 920
 DB |||||
 DB 1005 ATGATAACATCAAGCAGTGTCTTTGAACTTACTTTGAGGACTACAGGTACAACTGAAG 1064
 QY 921 AAGCTACAGATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAGCAAC 980
 DB |||||
 DB 1065 TCACAACAATCTGGGCCATGACTGAACATAAATAAATACTCAAGCATTTATGAAGAAAGTGC 1124
 QY 981 AAGAAGAGTTAGACAAGTATTTGTGAAATCGGAAAGCTTGAATCAATCAAGATTTTCATG 1040
 DB |||||
 DB 1125 AAGAAGAAATTAGAGGCTTTAGTGTAAAGAAAGATTTCTTTGGATGAAGATATATTCAAA 1184
 QY 1041 ATTTGAAATTTCTCAAGTTTAGTGGTTTAAAGAAACTCTAAGATTACA---TCCTCCGGTTG 1097
 DB |||||
 DB 1185 AGTTTTCCTACTTTAAGGCGAGTGAATAAGAGACACTTAGATTGCACTTTACCAGCACCAC 1244
 QY 1098 TCTTGAATTCGAGGAGGTGTAGAGAAACAAACAGAAATTTGATGATATGAATTTTCAATCCGA 1157
 DB |||||
 DB 1245 TACTTATACCAAGAGAGAAACAAATAAAAAATGCAATGTTAGATGGCTACGAAATTTCCAGCCA 1304
 QY 1158 ACACCTCGAATTTGTTGTAATGCTTGGCGATAGGAGAGATCCTTAATATCTTCTCGGAAAC 1217
 DB |||||
 DB 1305 AGACATTTATATGTAATGCTTGGGCAATTCATAGAGATCTTAAGGCTTGGAAAGATC 1364
 QY 1218 CTGAAAGTTTAAACCCGAAAGGTTTAAAGATTGTGCAATTTGTGCAATTTATAAAGGAGACAT 1277

Db 1365 CAGAGAGTTTATACCTGAGAGGTTTTTAAATTTGGACATAGATCTTTATGGACAAGATT 1424
Qy 1278 TTGAACCTGGTACATTTGGTGCAGAAAAGAAATATGCTCGCATTAATCTTCCAGCTATTA 1337
Db 1425 TTGAGTTTATTCATTCGGTGTGCTGATGATTTGGCTGATGATGATGATGATGATGATG 1484
Qy 1338 CCAATTTGGAGTATGATTAATAATTAATTAATTAATTAATTAATTTTAAATTTGGAACTGGCCGATG 1397
Db 1485 CTGCATTGGATCTTATCTTACTTGTAAATCTTCTTATTTCTTGTAGCTGGAAATTTGCCCAAG 1544
Qy 1398 GAATTACACCTCAACACCTTGATA 1421
Db 1545 GAATGAAAAGGAAGACATTGATA 1568

RESULT 4
US-10-953-349-13595
; Sequence 13595, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 2004-09-30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13595
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-13595

Query Match 12.9%; Score 224.2; DB 6; Length 1915;
Best Local Similarity 51.2%; Pred. No. 38-50;
Matches 762; Conservative 0; Mismatches 668; Indels 57; Gaps 8;

Qy 48 TAATAAGTTTCTCGCTGTTGTTTAAATCTTAGTAGTAGTATCATGAGGTTGTGGAGAAACAGA 107
Db 65 TAAACAGTTTCTCTGCATCTTCTTATATCTTTCAGCTAATAAGAAACAAATACA 124
Qy 108 ATCCACCTCCAGGGCATGGAAGTTTCTTATCATAGGTAATCTTCTCATTTATTAATCA 167
Db 125 ATCTGCCACATCCCAACCAAGATACCCATAATCGCAATCTTCCACAGCTA----- 177
Qy 168 CTTCGTATAGCCATGAACGTTTATAGAGCCTTGGCTCAAAATTTATGACCTGTTATGA 227
Db 178 --GGCACACTGCCACACCGCTCTTTCATGCACTCTCACAAATATGGCCCTCTCATGA 235
Qy 228 GTCTTCAAAATGGCCAAAGTTTCAGCTGTTGTCATTTCTTTCAGCTGAAGCAGCAAGAGG 287
Db 236 TGTTCGAATTTGGTCAAAATTCACACCTAGTGTCTCATCAGCTGACGTGGCCAGAGAA 295
Qy 288 TTATGAAACTCAGGCTGATGCTTGGCCCAACGCGCTATCGCTTTGAGCAGCAGATG 347
Db 296 TAATCAAAACGCGATGATGTTGTTTCTTCCAAACCGCGCAACCTACAGCTGCTAAATCT 355
Qy 348 TGTTTTTATAATCGGAAGATGCTTTGCTTTCATATGAGATCACTGGAGGAGATGA 407
Db 356 TTGGTTATGATGCAAAAGATGCGCTTTCGTGTACTACCGCAAGAGTGGAGACAAAGA 415
Qy 408 AGAAAATTTGGATCTTGAATTTCTGAGTGCCAAAAGTTTCAATCTCCAGGTTTAATCC 467
Db 416 TAAAGACATGTAAGTTGAGCTTATGAGTCTGAAGAGTGCGGTTGTTTCATTCCAATTA 475
Qy 468 GAGAGA-----AGAAATGGAGGATGCCATCACAATCTCCGTTGCGAAAGCCGAT 518
Db 476 GACAAGAAGTTGTTACAGAGTTGGTTGAAGCTATAGGTGAACGCTGTGTTAGTGAAGAC 535
Qy 519 CTCGGTCAATATTCAAGATCAT-----TTATGGCATTAAT 558

Db 536 CATGTGTGAATCTGACTGAGATGCTGATGGCAGCATCGAACGACATTTGTGTCTAGATGTG 595
Qy 559 TTGATCATGATAAGAACATCCGTTGGTAAATTTGTAAGCAAAAGAAAGATTTGCTCAGTGT 618
Db 596 TTCTTGGACGGAAGTGTGATGATCATGTGTGTGTAGTGGCAGTAGCAGCTTTTGCAGCGT 655
Qy 619 TGCCGATGTCAGTCAATGAGGC-AGGACGAGTTTGGCACCAGCAGACGCTTTTCCGACGT 677
Db 656 TGGGAAGAAAGATTATGAGACTATATCGGCTTTTTCAGCGTGGGTGATTTCTTCCCTTGGT 715
Qy 678 GGAATTAATCTTCACTATATCATTTGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGA 737
Db 716 TGGTTGGTTGACTATCTGACTGGCTTAATTCAGAGATGAAACACGCTTTCTCGCAG 775
Qy 738 TTGACGATATATCTTGAAGAGATTCTTAATGAACACAAAGCCAAATGAAGCCTTTTGAAGCGG 797
Db 776 TAGATGCTTCTTGTGATGAGGTAATTCAGAAACACGAGAGCAGTAAACAAGAA--TG 832
Qy 798 ATAACTTAATGGATGTTCTTATTTGAATCTTCAAAAATGGAACGTTTCCAGTGCCAGTGA 857
Db 833 ATGACTTCTTGGGATACCTTCTTCAACTTCAAGAAATGTGGAGGCTTGACTTTCAGCTCG 892
Qy 858 CAAACGAAAGCATCAAGCATCCGTTTGTGCAAAATGTTTACTGCCGGGAGCGAAACAACTT 917
Db 893 ACCGAGATTAACCTCAAGCAATCTTAGTGACATGATTAATAGTGGGAGTGACACTACTT 952
Qy 918 CGAAAGCTTACAGAATGGGTAAATGGCAGAGCTGATGAAAAAATCCAACTGAACTAAGAAAAG 977
Db 953 CAACAACTCTAGAATGGACTTTTGGGAGTTCTTAGAAATCCAAATACCATGAAGAAAG 1012
Qy 978 CACAAGAAGAAGTTAGACAAGTATTG--TGAATGGGAAAGCTTGTGATGAATCAA 1031
Db 1013 CTCGAAGAAGAGGTGAAGAGTGTGGGAAATCAATTCCAAAGCAGTACTGGATGAAATTT 1072
Qy 1032 GATTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTTACATCTCT 1091
Db 1073 GTGTGAATCAATGAATCTTGTGAATGTGTAGTCAAGAAACTTTGAGATTTACATCCAC 1132
Qy 1092 CGGT---TGTCTTGAATCCGAGGAGTGTAGAGAAACCAACGAAATTTGATGGATATGAAA 1148
Db 1133 CCTTCTCTTGTGATGCTCGAGAGACATCATCAAGTGTAAACCTAAGAGGGTACGATA 1192
Qy 1149 TTCACTCCGACACATCGAATTTGTTGATGCTTGGCGGATAGGAGAGATCCTAATACTT 1208
Db 1193 TTCCCGCAAAACAAATGATTTATCAATGATGCGCGATCCAGAGGGATCTTGAATTA 1252
Qy 1209 GGTCCGACCTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGTGCAATTTGATTAAG 1268
Db 1253 GGGATGATCTCTGAAGATTTATTCGGAAGATTTGAAACCTAGCCAAGTTGATCTTAATG 1312
Qy 1269 GGACGACATTTGAACCTGGTACCAATTTGGTGCAGGAAAGAAATATGCTCTGGCATTAAT 1328
Db 1313 GACAAGATTTTCAATTAATTCGTTTGGTAATGGGAGAGGGGATGCGCTGCAATGTCTAT 1372
Qy 1329 CAGCTATTACCAATTTGGAGTATGCTATTAATAATCTATTAATCATTTTAAATTCGGGAC 1388
Db 1373 TTGACTTGTCTTCAACTGAGTATGTTCTGCTAATCTTTTGTATTTGTTTCAATGGAATA 1432
Qy 1389 TGGCCGA-----TGGAAATTACACCTCAAAACATTTGATATGACTGAAGCTATTTGGCGGTG 1442
Db 1433 TGTCCGAGTCTGGAGCTATATTGATGCACAACTTGACATGATGAGACAAATGACTCA 1492
Qy 1443 CTCTCAGAAAAAATAGATCTTAAGTTGATTTCTTATTTCCATATCAA 1489
Db 1493 CTGTGAGTAAGAAAGTACCACCTTCTATCTTTGAACCAAGAACCATATAAA 1539

RESULT 5
US-10-953-349-2239
; Sequence 2239, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.

Qy	173	GATCTAGGCCATGAACGTTTCTAGAGCCTTGGCTCAAAATTTATGACACCTGTTATGAGTCTT	233
Db	175	GAATTAACAACATCATGCTCTATGGAAGCTCTCAAAGAAGTATGCGCTTGATGCTTTTG	234
Qy	233	CAAAATTCGCCAAGTTTCAGCTGTTGTTCATTTCTTCAGCTGAAGCAGACCAAGAGGTTATG	292
Db	235	AAGCTTGGAAAGTCCCACTCATCTCTTCATCAGAAAAGCAGCAAAACAAGCTCTA	294
Qy	293	AAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACACAGATTGTGTTT	352
Db	295	AGAGACTATGATCTCCATGTTGTAGCGTCTCTTCTAGCAGGAGGAAGAGAGCTCTCT	354
Qy	353	TATTAATCGGAAAGATGCTTGTGTTGCTTCATATCGGAGATCACTCGAGCGCAGATGAAGAA	412
Db	355	TACAACAATCTAGATATGCTCTCATCTCTTATAATGAATATTGGAAAGGAACTAAGGAAG	414
Qy	413	ATTTTGGATATCTGAAATTTCTGAGTGCACAAAAGTTCAATCCTCCAGGTTAAATCCGAGAG	472
Db	415	CTCTGCAGTCAAGAACTCTTTAGTGTCTAATAAAATTCATTCGATTCAACCTATTAAAGAC	474
Qy	473	GAAGAAATGGAGGATGCCATCACATTCCTC-----CGTTCGAAAGCCGGAATCTCCGGTC	526
Db	475	GAGGAGTCAAGAAAGTTATCGATTCAATCGCGGAATCATCTTCTCTAAAGAATCCGGTT	534
Qy	527	AATATTACAAGATCAATTTATGGCATTTAATTTTCGATCATGATAAGAACATCCGTTGGT	586
Db	535	AACCTTGACAAAGCGTTTCTTGCTTTAACTACAAGTGTAGTATGCAAGGCAGACATTTGGT	594
Qy	587	AATTGTAAGCAAAAAGAAAGATTGCTGAGTGTTCGCGATGCAGTCAATAGGCAGCGACG	646
Db	595	GTGAGTTTTGAGGGAAGTGTCTCAACAGTGATA--GATTCAATAAGTTAGTCCGAGATA	652
Qy	647	AGTTTGGCACCGCAGAGCGTTTTCGACGTGGAAATTACTTCACTATATCATTTGGAGCT	706
Db	653	CTTTCGAGATGTTGGGAAGCTTCTCGCTCAGATTTTATTCGCTATGTGCGGATGGATCA	712
Qy	707	GAGTCAAAACCCAGGCGCTTTGCA-----TCAGGAGATTGACGATA	746
Db	713	TCGATAGTTCAATGGTTTACAGGGTGGAGAAAGAAAGCTTTCGAGATCTTGATCGGT	772
Qy	747	TACTTGAAGAGATCTTTTAATGAACACAAAGCCATAAGCCCTTTTGAAGCGGATAACTTAA	806
Db	773	TCATATGAACAAATTTTTGATCTGCATTAAGGAAGAGGAAGTAGGAAGTGAAGATTTTGA	832
Qy	807	TGATGTTCTATTGAATCTTCAAAAAATGGAAACGTTCCAGTGCCAG-----TGNCAA	860
Db	833	TGATGTGCTCTTGAGGTGGAGAAAGAGAAATTTGTTGTGGAAATGGCAAGCTCAAA	892
Qy	861	ACGAAAGCATCAAGCATCGGTTTTGCAAAATGTTTACTGCGGAGGCGAAACAACCTTCGA	920
Db	893	GAATCATATCAAGCAATTTTGATGAACATCTTTTAGGAGGAAATCGATCTTCTCGAA	952
Qy	921	AAGCTACAGAATGGGTAATGGCAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCAC	980
Db	953	TAACAATGACATGGGCAATGGCGAACTTGCTAAAAACCTTAGAGTGATCAAGAAAGTTC	1012

Db 147 TACATCAGATTGGAGAAATGCTCTCAAGGTCACATTTCAACAT-----CTCGCGGAAA 197
Qy 210 TTTATGGACCTGTTATGATGCTTCAAAATTTGGCAAGTTTTCAGCTGTTGTCATTTCTTCAG 269
Db 198 GATCGGACCTGTGATGCTTCTTCACTTTGGTGTTCCTTAACTGTGCTCATCGA 257
Qy 270 CTGAAGACGCAAGAGGTTATGAAAACTCAGGCTGATGCCCTTCGCCCAAGCCCTATCG 329
Db 258 GAGAAGCCGCTCAAGAAGTGTAGAACTCATGACCTAGACTGTTGTCAGCAGGCTTAAGC 317
Qy 330 TCTTGGACGACAGATGTTGTTTTATATCGGAAGATGTCCTGTTGCTTCATATGAG 389
Db 318 TTGTCGGACAAAGTTACTCTCGCGGATTTTAAAGATATCGGTTTTACGCCATACGGTA 377
Qy 390 ATCACTGAGGAGCAGATGAAGAAAATTTGGATACTTGAATTTCTGAGTCGCAAAAAGTTC 449
Db 378 ACGAGTGAAGGCGGGGTAAAGTTGGCTGCGTGTGAGCTTTCTGTTGAAAAGGTTTC 437
Qy 450 AATCCTCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTCCT-----CC 503
Db 438 AGTCTCTTAGGCATATCCGAGAGGAAGAAATGTAACCTTCTGCTCAAGCAACTGTCCGAAT 497
Qy 504 GTTCGAAAGCCGGATCTCGGTCAATATTAACAAAGATCATTTATGGCATTTAATTCGA 563
Db 498 CTGCGGTTGATCGCTCTCGGTCGATTTGAGCAAAATCCCTTTTCTGGCTAACCGCTAGTA 557
Qy 564 TCATGATGAACATCCGTTG-----TAATTGAAGC 596
Db 558 TCCTTTTAGAGTTGCCTTAGCAGAGAAATTTTACGAGAGCGATTTTATCGATAAGAAA 617
Qy 597 AAAAAAGAAAGATTGCTGAGTGTTCGGATGCAATGAGGACGACGAGTTTGGCA 656
Db 618 AGATCGAGAGCTCGTGTTCGAAGCTGAGACTGCCCTAGCAAGTTTCACTGTTCTGATT 677
Qy 657 CCGCAGACGCTTTCCGAGCTGGAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAAC 716
Db 678 TCTTCCCTGTTGCGGAGCTTGATGGCTGTTGATTGGTTTTCCGGCAACAAAGAGAC 737
Qy 717 CCAGGCTTTTGCATCAGGAGATTGACGATATCTTGAAGAGATTCTTAATGAACAAAG 776
Db 738 TCACGATGTTTTTTACAGCTCGATGCTCTGTTTCAACATGTCATAGATGATCATTTAA 797
Qy 777 CCAATAAGCCTTTGAAGCGGATAACTTAATGGATGTTCTATG-----AATCTTC 827
Db 798 ATCTCGAAGATCAAAAGAGCAGGAACATCATCGAATTCATGTTGGATGTGATTCATA 857
Qy 828 AAAAAATGGAACGTTCCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTC 887
Db 858 AACAAAGGAGAGACAGTTCCTTAGAGCTCACAAATAGATCATATCAAGGGGTTTCTCGGA 917
Qy 888 AATGTTTACTCGGGAGCGGAAAACAACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGC 947
Db 918 ATATATTTCTCAGGGATAGACACAGGGGCCATCACCATGATATGGGCAGTGAACGAGC 977
Qy 948 TGATGAAAAATCCAAGTAAAGAAAAGCAAGAAAGATTAGACAAAGTATTG--- 1004
Db 978 TCGTTAAAAAACCGAACTGATAAGAAAAGTTCAAGGCGATATCCGAGAACAACTTGGCA 1037
Qy 1005 GTGAAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGG 1064
Db 1038 GCAATAAGGAGAGAAATCACCGAGGAAGATATCGAGAAAGTTCCTTACTTTGAAGATGGTAA 1097
Qy 1065 TTAAGAAATCTTAAGATTACATCTCTCGGTTG---TCTTGATTCGAGGGAGTGTAGAG 1121
Db 1098 TCAAGAAACATTCAGTTTACCCAGCAGCTCCTCTTATATCTTCAAGGGAAACATAGG 1157
Qy 1122 AAACAACACGAATTGATGATATGAATTCATCCGAACACTCGAATTTGTTGTAATGCTT 1181
Db 1158 CTCACATCAAGTTCAAGGGTATGATTTCTCTCCCAAGAGGAGGATCTTGGTCAATGTTT 1217
Qy 1182 GGGCGATAGGAGAGATCTTAATACTTGTGTCGGAACTCGAAAGTTTAACCCAGAAAGGT 1241

Db 1218 CGGCAATAGGAGAGATCCCAAACTCTGGACAAACCCGAAAGAGTTTACCCCTGAGAGT 1277
Qy 1242 TTAAAGATTGTGCAATTTGATTATAAAGGACGACATTTGAACTGGTACCATTTGGTGGAG 1301
Db 1278 TTATGGATAGCTTTGTTGATTATAGGGACAAACATTACGAGCTCTTACCAATTTGGGTCG 1337
Qy 1302 GAAAAAGATATGCTCTGGCAATTACTTCAAGCTATTAACAAATTTGGAGTATGTCATTATA 1361
Db 1338 GTCGAAGGATATGCTCCGGGATGCCAATGGGGATTGCTGCCGTCGAATTTGGACTCTTGA 1397
Qy 1362 ATCTATTATATCATTTTAAATTTGGGAACTGGCCGATGGAATTTACACCTCAAAACACTTGATA 1421
Db 1398 ACTTACTTTACTTCTTCAATTTGGAAGTTGCTGATGGATGCACATATAAGATATCGATA 1457
Qy 1422 TGACTGAAGCTATTGGCGGTCTCTCAGGAAAAAATAGATCTTAAAGTTGATTCCTATTC 1481
Db 1458 CTGAAGAAGCTGGTACTCTTACAATAGTCAAGAAAGTACCTCTCAAGCTCGTTCAGTTTC 1517
Qy 1482 CATATCA 1488
Db 1518 GAGTTCA 1524

RESULT 7
US-10-953-349-17424
; Sequence 17424, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17424
; LENGTH: 1958
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-17424

Query Match 10.0%; Score 172.8; DB 6; Length 1958;
Best Local Similarity 57.3%; Pred. No. 1.2e-36;
Matches 353; Conservative 0; Mismatches 257; Indels 6; Gaps 2;

Qy 877 ATCCGTTTTCGAAATGTTTACTGCGGAGCGGAAACAACTTCGAAAGCTTACGAATGGGT 936
Db 1185 ATATTTTCAGGATATGTTTGTAGCAGGAAGTGAACAACCGCATCAGCACTAGAAATGGGC 1244
Qy 937 AATGCGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAGCAAGAAGAAGTTAGACA 996
Db 1245 TATGCGAGCTCATGAAAAATCCAAATGAAATTTGAAGAAAGCCAGGATGAGGTGAGAAA 1304
Qy 997 AGTATTTGCTGAAATGGGAAAAGTTGATGAATCAAGATTTCATGATTTGAAATTTCTTCAA 1056
Db 1305 ATTTGAGGGATAAATCAAAAGTAGAGAAAATGACATCAATCAATGGACTACATGAA 1364
Qy 1057 GTTAGTGTGTTAAAGAACTCTAAGATTACATCTCCTCGG---TTGTCCTTTGATTTCCGAGGGA 1113
Db 1365 ATGTGTAATCAAGAAACTCTAAGGTTACATCCACCTGCTCCTCTCTTGGCTCCTAGAGA 1424
Qy 1114 GTGTAGAAAACAAACGCAATTTGATGGATATGAATAATTCATCCGAACACTCGAATTTGTTGT 1173
Db 1425 GACAGCATCTAGTGTGAAACTGGGAGGGTATGATATTCAGCAAAAACACTCGGTATATGT 1484
Qy 1174 GAATGCTTTGGCGATAGGAGAGATCCCTAATACTTTGGTCGGAACCTGGAAAGTTTAAACCC 1233
Db 1485 CAATGCATGGGCAATTCAGAGGGACCCCTGAGTTTGGGAAAGACCTGAAGAGTTCAATCC 1544
Qy 1234 AGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGGACGACATTTGAACCTGGTACCATT 1293
Db 1545 AGAGAGACATGATATAACAGCCGGGTTTCATTTAATGCCAAGACTTGCATTTTATTACATT 1604

QY 1294 TGGTGCAGGAAAAAATATGCTCGCATTACTTCAGCTATTACCAATTTGGAGTATGT 1353
 |||||
 Db 1605 TGGTTTGGAGAGGGCATGCCCGAATGAGTTTGGCTTCTCTTGGATATAT 1664
 QY 1354 CATTATAAATCTATTATATCAATTTTAAATGGGAACGGCGATGGAATTAACACCT---CA 1410
 |||||
 Db 1665 TCTGGCCAACTTCTTTTATGGTTCAATTGGAAGCTGCCCTGCAACTCACACATCTGGACA 1724
 QY 1411 AACACTTGTATGACTGAAGCTATTGGCGTGCTCTCAGGAAAAAATAGATCTTAAGTT 1470
 |||||
 Db 1725 AGACATAGACATGAGTGAACATATGGCTAGTTTACCTACAAGAAAGAGCACTTTCATCT 1784
 QY 1471 GATTCCTATTCCATAT 1486
 |||||
 Db 1785 TAAACCAATCCCTTT 1800

RESULT 8

US-10-953-349-4847
 ; Sequence 4847, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 4847
 ; LENGTH: 1586
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-953-349-4847

Query Match 9.6%; Score 166; DB 6; Length 1586;
 Best Local Similarity 49.1%; Pred. No. 6.9e-35;
 Matches 705; Conservative 0; Mismatches 680; Indels 51; Gaps 8;

QY 2 CATATAAGGAAATGGAGCAGAAAAATCTCTTTTCCGAGCATTTTAAATAAGTTTCTG 61
 |||||
 Db 17 CAAAAACAAAAACAGAGCCAGGAATAATGTCACTTCTCTCTCTCTCTCTCTCTCTCT 76
 QY 62 CTTGTTTAACTTAGTAGTAGTCATGAGTTGTGGAAGAAACAGATCCACCTCCAGG 121
 |||||
 Db 77 CCCCTTATCTTAATCTTCTTGAATAATCTCAAAACCATCGAAATGGAAGCTTCTCCAGGC 136
 QY 122 CCATGGAAGTTTCTTATCATAGTAACTCTCTCTTATTTACTCACTTCTGATCTAGGC 181
 |||||
 Db 137 CCAAGAGAGCTTCGATCATCGGAA-----CTTACACCAACGCCGGGAATTACATCC 189
 QY 182 CATGAACGTTTATAGCCTTGCTCAAAATTAAGACCTGTTATGAGTCTTCAAAATGGC 241
 |||||
 Db 190 CAGGAACAGTCGGAATCTTTT---CCGAAAAGTACGGACCAATCGTGTCTCCGATACGGA 247
 QY 242 CAAGTTTCAGCTGTTGTCATTTCTTCAGCTGAAGCAGCAAGAGAGTTATGAAAACCTAG 301
 |||||
 Db 248 TTCGTCCTCCGCTGGTGTGATCTCGTCAAAAGAGCAGAGGAAAGTTCTCAAGACCCAC 307
 QY 302 GCTGATGCTTCCGCCCAACGCCCTATCGTCTTGGAGCGACAGATGTGTTTTATAATCGG 361
 |||||
 Db 308 GATCTGAGTGTGTAGCCGACAGAGAGCGTTGGGACCAAGCAATCTCTTACACTTT 367
 QY 362 AAAGATGCTCTGTTGCTTTCATATGAGATCACTGGAGCAGATGAAGAAAAATTTGGATA 421
 |||||
 Db 368 AAAGACATCGGATTCGACCGTACGTTGAAGATTGGAGAACGATGCGGAAGCTCTCGGTG 427
 QY 422 CTTGAATTTCTGAGTCCCAAAAGTTCAATCTCCAGGTTAATCCGAGAGCAAGAAATG 481
 |||||
 Db 428 GTCGAGCTCTTCAGCTCGAAAAAGCTTCAATCTTTCAGGTATATCAGAGAGGAAGAAC 487

QY 482 GAGGATGCCATCACATTTCTCGTTCGAAAGCG-----GATCTCGGTCAATATTACA 535
 |||||
 Db 488 GACTTGTGTCAAGAACTCTCTGATTTAGCTTCGAGACGATCTTTGGTGAATCTTGAG 547
 QY 536 AAGATCAITTTATGGCAATTATAATTTTCGATCATGATAAGAACATCCCGTTGGTAAATTTGAAG 595
 |||||
 Db 548 AAAACCCCTTTTACTTTTAGTCGGAAGTATAGTGTAGGATAGGGTTGGGATAAATCTC 607
 QY 596 C-----AAAAAGAAAGATTCTCGAGTGTTCGCCGATTCGCGATCAATGAG 637
 |||||
 Db 608 CFTGAGTGTGAGTTTCGTTGATGAAGATAGCATCGATGATCTCTGTGCACAGTCTGAGAT 667
 QY 638 CGAGCCAGCAGTTTGGCACCGCAGACGCTTTTCC-----GACGTGGAAATTTACTTCCAC 691
 |||||
 Db 668 GTCATAAGGAATCTATCTCTGATTTCTTCCCGGATTAAATGGTAGGCTCATCGAG 727
 QY 692 TATATCATTGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTT 751
 |||||
 Db 728 TGGATCTTCAGTGAGAGGAAGAGATTGAATAGACTTTTACTCGGAAGTAGACACTTTCTTT 787
 QY 752 GAAGAGATTCTTAATGAAC-----ACAAAGCCAATAAGCCTTTTGAAGCGGATAACTTA 805
 |||||
 Db 788 CAGAATATTTCTTGTATGATCATCTCAAGCCTGGAAGAGAGAGCTCTGATATCATCGATGTG 847
 QY 806 ATGGATGTTTCTATTGAATCTTCAAAAAAATGAAACGTTTCCAGTGCCAGTGACAAACGAA 865
 |||||
 Db 848 ATGATTGATATGATGAAGAGCAAGAAAGAGGAGACTCTTTCAAGTTCCACCTGAT 907
 QY 866 AGCATCAAAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGAACAACATTTTCGAAAGCT 925
 |||||
 Db 908 CATCTCAAAGGGATGATCTCGGACATATTTCTAGCAGGAGTTTGGAAACAAGCTCCACTACA 967
 QY 926 ACAGATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACATAGAAAAAGCAAGAA 985
 |||||
 Db 968 TTGATATGGGCGATGATCCGAGCTGATCAGAAACCCGAGAGTGTATGAAGAAAGTTCAAGAC 1027
 QY 986 GAAGTTTAGACAAGTATTTGGTCAAAAT---GGCAAAAGTTGATGAATCAAGATTTTCATGAT 1042
 |||||
 Db 1028 GAGATTCGAGACAACACTTGGGACAGAAAGGAGAGATCACAGAGAAGATTTAAACCAG 1087
 QY 1043 TTGAAATTTTCAAGTTAGTGTGTTAAAGAAACTCTAAGATTACATCTCTCCGGTT---GTC 1099
 |||||
 Db 1088 CTTCACTACTTCAAGCTCATGCTGCTCAAGAGAGATATTCAAGTTTACATCCAGCAGCTCCACTT 1147
 QY 1100 TTGATTCGAGGAGGTGTAGAAAAACAACAACAATTTGATGGATATGAATTTATCCGAAC 1159
 |||||
 Db 1148 TTGCTCCCAAGAGAGACATTTGCTCATGTCAAGATCCAGGCTACGATATTCTGCTAAA 1207
 QY 1160 ACTCGAATTTGTTGAAATGCTTGGGCGATAGGAAGAGATCCCTAATCTTGGTGGAACTT 1219
 |||||
 Db 1208 ACACAGATCATGATCAACGCTTACCGGATTGACGCTGATCCAAACTATGGACAAACCTT 1267
 QY 1220 GGAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTATAAGGGAGCGACATTT 1279
 |||||
 Db 1268 GATGAGTTTAAACCCCTGATAGGTTCTTGACAGCTCCATAGATTACAGAGGACTGAACTTT 1327
 QY 1280 GAACTGGTACCATTTGGTGCAGAAAAAGAAATATCTCTCGCATTTACTTCAGCTATTACC 1339
 |||||
 Db 1328 GAGCTATTACCGTTTGGATCTGGTAGGAGGATATGTCAGGAGTGAATATGGGATCGCC 1387
 QY 1340 AATTGGAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGAACTGGCCGA 1395
 |||||
 Db 1388 ATTGTTGAATTGGGACTATTGNAITTTGCTTACTTCTTCGACTGGGGGTACAGA 1443

RESULT 9

US-10-953-349-13112
 ; Sequence 13112, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; TITLE OF INVENTION: ENCODED THERBY


```
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13112
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (474)..(474)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-13112

      Query Match      9.2%; Score 159; DB 6; Length 1659;
      Best Local Similarity 56.4%; Pred. No. 5e-33;
      Matches 341; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY      807 TGGATGTTCTATTGTAATCTTCAAAAAAATGGAACGTTCCAGTGCCAGTGACAAACGAAA 866
DB      852 TTGATGTTTGTGTCATATCTGAAGATGAGAGTTCTGMAATTAAATTAACAAAGAAA 911

QY      867 GCATCAAGCATCCGTTTTCGAATGTTTACTGCCGGGAGCGAACAACCTTCGAAGCTA 926
DB      912 ACAATTAAAGGCCCTTCATCTTGGACATATTAAATTGCTGGGACTGACACCTCAGCTGTAACGA 971

QY      927 CAGAATGGTATGTCAGAGCTGATGAAAAATCCAACTGMACTAAGAAAAAGCACAAAG 986
DB      972 TCGAATGGGCTATGCGAGAGTTAATCAACATCCAGGTGTGTGGAGAGGCGACAGCAAG 1031

QY      987 AAGTTAGACAAGTATTGTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGA 1046
DB      1032 NAATGGATGCAGTGGTTGGAAAGAGTAGAATAGTAGAAGATCAGATATTGCCAACCTTC 1091

QY      1047 AATCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCCTCGGTTGTCTTGATTC 1106
DB      1092 CTTACTTGCAGGCCATTGTTAGAGAAACATTAAGACTTCCACCAGCTGGTCCATTGCTTT 1151

QY      1107 CGAGGGAGTGTAGAGAAACACACGAATTTGATGGATATGAATTCATCCGAACACTCGAA 1166
DB      1152 TTAGAGAGTCATCTAGAGAGCTGTGGTTTGGGTATGATATTCCAGCAAGACTCGAT 1211

QY      1167 TTGTTGTGAATCTTGGGCGATAGGAAGAGATCCTTAATCTTTGGTCGGAACCTGGAAGT 1226
DB      1212 TATTTGTCAATCTTTGGGCTATTGGTAGGACCCCAATCCTCTGGGAGAACCTCTTGAGT 1271

QY      1227 TTAACCCAGAAAGTTTAAAGATTGTG-----CAATTGATTATAAGGGACGACATT 1278
DB      1272 TTAGGCCAGAGAGTTTGTGAAAATGGGAAGAGTCAATTGGATGTTAGGGGACAACTT 1331

QY      1279 -TGAACCTGTACCATTTGGTCAGGAAAAGAAATATGCTCGGCATTACTTCAGCTATTA 1337
DB      1332 ATCATCTACTTCGTTCCGTTAGTAGGAGAGCATGTCCTGCTACTTCTTTGGCATTCG 1391

QY      1338 CCAATTTGGAGTATGTCATTATAAATCTAATATATCATTTTAATTGGGAACCTGGCCGATG 1397
DB      1392 AAGTTGTGCATGTAATTTGGCAGTCTTAATTCAGTGTTCCTCAATGAAGAGTTGACTGTG 1451

QY      1398 GAATT 1402
DB      1452 ACAAT 1456

RESULT 10
US-10-953-349-17280
; Sequence 17280, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
```

```
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17280
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-953-349-17280

      Query Match      9.1%; Score 158.4; DB 6; Length 1658;
      Best Local Similarity 49.6%; Pred. No. 7.2e-33;
      Matches 643; Conservative 0; Mismatches 596; Indels 57; Gaps 7;

QY      164 CTCACCTTCTGATCTAGGCCATGAACGTTTATAGACCTTGGCTCAAAATTTATGGACCTGTT 223
DB      164 CTTCTTCTCCACACACCTCACCAAGATTTTCAACAGCTCTCACTCCGCTATGGACCCATA 223

QY      224 ATGATCTTCAAAATGGCCAAGTTTTCAGCTGTTTCTCATTTCTTTCAGCTGGAAGCAGCCAAA 283
DB      224 ATACACCTTTTCTGTTTTCAGTCCCTGTGTGGTGGCTTCCACAGCAGAGAGCCGCCAAA 283

QY      284 GAGGTTATGAAACTCAGGCTGATGCCCTTCGCCCAACGCC---CTATCGTCTTTGGACGCA 340
DB      284 GAGTTCTCAAACTCAGAACCCGCTTCTCCAACCCGCCCTAACACTGTGCGCGTC 343

QY      341 CAGATTGTTTTTATAATCGGAAAGATGTCCTTGTGTTTGTTCATATGGAGATCACTGGAGG 400
DB      344 GAAACCTTAACCTACGGTTTCCAAAGACTTCTTGTTCGACCCCTACGACCCCTATTGGAAG 403

QY      401 CAGATGAAGAAATTTGGATACCTTGAATTTCTGAGTGCCAAAGAAAGTTCAATCCTCCAGG 460
DB      404 TTCATGAAGAAACTCTGCACTGCCAACTCTCGGTGGCCACATGCTCGACCAAGTTTCTT 463

QY      461 TTAATCCGAGAGGAAGAAATGGAGATGCCATCACATTCCTCCGTTCCGAAAGCCGATCT 520
DB      464 CCCGTGAGACAGCAGAGACAAAGAAATTCATCAACCGTGTCTCCAAAGGGTATTTCT 523

QY      521 CCGG-----TCAATATTACAAGATCATTTATGGCATT 553
DB      524 GGTGAGGCGGTGGATTTTGGGGGAGAGTTTCATACGCTCTCGAACCAATCGTGTGAGA 583

QY      554 ATAATTTTCATCATGATGAAGAACATCCGTTTGGTAAATTTGAAGCAAGAAAGAA-----AAGA 607
DB      584 ATGATCGTGAAGTACGACGAGTACTACTGAGACGAGAACAGAAAGTTGAAGAGATGAGGAAG 643

QY      608 TTGCTGAGTGTTCGCGATGTCAGTCAATGAGGCGACGAGTTTTTGGCACCCGACACGCT 667
DB      644 CTGTTGAAGGATGCGCAGAGCTCTCGGGGAAAGTTTCAACATATCGACTTCGTTTCGTTTC 703

QY      668 TTTCCGACGCTGGAATTTACTTCACTATATCATTTGAGCTGAGTCAAAACCCAGGCGTTTG 727
DB      704 TTGAAGCGCTTTGATTTTGCAGGGTTCAACAGAGGCTCGAGAGAGATTTCGGGACTGCTTT 763

QY      728 CATCAGGAGATTGACGATAT---ACTTGAAGAGATTCTTAATGAACACAAAGCCCAATAA 783
DB      764 GACACCGTTTGGACAGAGATCATTAAGCAACGTAAGAGGAAGAGGAACAGATGAA 823

QY      784 GCCTTTTGAAGCGGATAACTTAA-----TGGATGTTTCTATTGAATCTTCAAAAAAAT 835
DB      824 ACCGTTGGAACAGAGATTTAAGGATATGCTTGATGTTTGTGTTTGGACATATCTGAAGAT 883

QY      836 GGAACGTTCCAGTGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTGTGCAATGTTT 895
DB      884 GAGAGTTCTGAAATTTAAATTAACAAAGAAACATTTAAGGCCCTTCATCTTGGACATATTA 943

QY      896 ACTCCGCGGAGCGAAACCACTTCGAAAGCTACAGATGGGTAAATGGCAGAGCTGATGAAA 955
DB      944 ATTGCTGGGACTGACACCTCAGCTGTAAGTGAATGGGCTATGGCAGAGTTAATCAAC 1003

QY      956 AATCAACTGAACTAAGAAAAAGCAGAAAGAAAGTTAGACAAGTATTTGGTGAATGGGA 1015
DB      1004 AATCCAGGTGTGTTGGAGAGGAGCAAGAAATGGATGCGAGTGGTTGGAAAGAGTAGA 1063
```



```
QY 1016 AAAGTTGATGAATCAAGATTTTCATGATTTTGAAATTTCTCAAGTTAGTGGTTAAAGAACT 1075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1064 ATAGTAGAAGATCAGATATGCGAACCTTCTTACTTGCAGGCAATGTTAGAGAAACA 1123
QY 1076 CTAAGATTACATCCTCGGTTGTCCTTGAATCCGAGGAGTGTAGAGAAACAACACGAAT 1135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1124 TTAAGACTTCACCCAGCTGGTCCATTTGCTTTTATAGAGAGTCACTAGAGAGCTGTGGTT 1183
QY 1136 GATGGATATGAATTCATCCGACACTCGAATTTGTTGTAATGCTTTGGCGATAGGAAGA 1195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1184 TGTGGTATGATATCCAGCAAGACTCGAATTTTGTCAATTTTGTCAATGTTTGGGCTATTTGAGG 1243
QY 1196 GATCCTTAATCTTGGTCGGAACCTCGGAAGTTTAAACCCAGAAAGTTTAAAGATTCTG-- 1253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1244 GACCCCAATCACTGGGAGAACCTCTTGAGTTTAGCCAGAGAGTTTGTGAAATGGG 1303
QY 1254 -----CAATTGATTAAGGACGACATTT-GAACTGGTACCATTTGGTGCAGGAAAA 1306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1304 AAGAGTCAATTTGGATGTTAGGGGACAACTTATCATCTACTTCGTTCCGTTAGTGAAGA 1363
QY 1307 AGAATATGCTGGCATTTACTTCAGCTATTACCAATTTGGAGTATGTCATTATAATCTA 1366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1364 AGAGCATGCTGGTACTTCTTTGGCAATTCGAAGTTGTGCATGTGAATTTGGCAGTTCTA 1423
QY 1367 TTATATCAATTTAAATGGGAACCTGGCGGATGGAATT 1402
Db 1424 ATTCAGTGTTCCAATGGAGGTTGACTGTGCAAT 1459
```

```
RESULT 11
US-10-953-349-10369
; Sequence 10369, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10369
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10369
```

```
Query Match      8.9%; Score 153.8; DB 6; Length 1722;
Best Local Similarity 53.2%; Pred. No. 1.2e-31;
Matches 326; Conservative 0; Mismatches 287; Indels 0; Gaps 0;
```

```
QY 853 AOTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTCGCGGAGCGAAAC 912
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 953 ATTAACAATGAATGATCTCAACACATCTGCTTCGATGTGTTTGTGTCGCACAGAC 1012
QY 913 RACTTCGAAAGCTACAGATGGTAAATGCGAGCTGATGAAAATCCAACTGAACATAAG 972
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1013 AAACCTCTAGTCAATGGAATGGCAATGACCGAGTTATTCCTGAGCAGGAAGATGGT 1072
QY 973 AAAAGCACAAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAG 1032
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1073 CAAGCAGACAGTGAATACGCGAAGTATTTGGTCAAAACGGTTTCGTTCAAGATCTGA 1132
QY 1033 ATTTCAATGTTGAAATCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCCTCC 1092
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1133 TATCCGAGTCTGCCTTACTTACAAGCAATTTGAAAGAGACTCTTCGTTTCATCCTCG 1192
QY 1093 GGTGTGCTTGATCCGAGGAGTGTAGAGAAACAACAGCAATTTGATGGATATGAAATCA 1152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1193 AGCTCCTTTGATCCCTAGAAAAATCAGAATCCGATGTTCAATATTTGGGTTTCTTCTTCC 1252
```

```
QY 1153 TCCGAACACTCGAATGTTGTAATCTTTGGCGATAGGAAGAGATCTCTAATACTTTGGTC 1212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1253 TATAACACCCAGGTTTGTGTGAACGTATGGCGATAGACGAGACGCGAGCGTGTGGGA 1312
QY 1213 GGAACCTGGAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTATAAGGGAC 1272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1313 AAATCCAATGAAGTTTCGAGCCAGAGAGGTTCTTGTGTACGAGAAACCGATGTAAGAGCAG 1372
QY 1273 GACATTTGAACCTGGTACCAATTTGGTCAGGAAAAAGAAATATGTCTCGGCATTTACTTCAGC 1332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1373 AGATTTTCGAGTTGATACCTGTTTGGATCAGGAAAGAGATGTGTCGCGAATCTCGATGGC 1432
QY 1333 TATTACCAATTTGGAGTATGTCAATATAAATCTTATATATCAATTTTAAATGGGAACCTGGC 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1433 TCTTAAGACAATGCATATGCTGCTTGCCTCTCTTCTTATTCCTTTGACTGGAGCTTCA 1492
QY 1393 CGATGGAATTTACACCTCAACACTTGTATGATGAAAGCTATTTGGCGGTGCTCTCAGGAA 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1493 AAACGGTGTGTCCTCCCGGAAACATTCACATGACGAGACTTTCCGTTCTTACCTTACACAA 1552
QY 1453 AAAAATAGATCTT 1465
Db 1553 GGCCAAATCTCTT 1565
```

```
RESULT 12
US-10-953-349-21323
; Sequence 21323, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21323
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-21323
```

```
Query Match      8.8%; Score 152.2; DB 6; Length 1660;
Best Local Similarity 54.8%; Pred. No. 3.2e-31;
Matches 345; Conservative 0; Mismatches 278; Indels 6; Gaps 2;
```

```
QY 764 AATGAACACAAAGCCAATAAGCCTTTTGAAGCGGATACTTAATGGATGTTCTTATTGAAT 823
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 823 AGTGTGTAAGAAAGATAATCACAACAAGAGTCAAGATATTTATGATATTTCTCTCCAG 882
QY 824 CTTCAAAAAAATGGAAGCTTCCAGTGCAGTGACAAACGAAAGCATCAAGCATCCGTT 883
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 883 CTACTTGTATGATCGTTTCTTCACTTTGATCTCACTCTCGACACATAAAAGCCGTGCTC 942
QY 884 TTGCAAAATGTTTACTCGCGGAGCGAAACAACCTCGAAGCTACAGAATGGGTAAATGGCA 943
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 943 ATGAACATCTTTTATAGCAGGAACAGACCCGAGTTCGCGACAATAGTTTGGGCAATGAAT 1002
QY 944 GAGCTGATGAAAAATCCCACTGAACATAAGAAAAAGCACAAAGAGTTAGACAAGTATTT 1003
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1003 GCATGTTTGAAGAAATCCCAATGTGATGAGCAAGTTTCAAGGAGAGTGAAGAAATCTATTC 1062
QY 1004 GGTGAAATGGGAAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTG 1063
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1063 GGTGCAAAAGATTTTCAATAAACGAAAGATGATGTGCAAGAGCTTCTTATCTCAAGAGCAGTG 1122
QY 1064 GTTAAAGAACTCTAAGATTA---CATCTCCGGTTGCTTCTGATTTCCGAGGAGGTGAGA 1120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1123 GTGAAGGAGACATTAAGATTTATCCACCTTCACCACACTTTTCCCAAGGGGTAACAATG 1182
QY 1121 GAAACACACGAATTTGATGGATATGAAATTCATCCGAACACTCGAATTTGTTGTGAATGCT 1180
```

Db 1183 GAAACATGCAACATAGAGGGTACGAAATTCAGAGCAAAACATATATAGTCATGTTAATGCA 1242
Qy 1181 TGGCGGATAGGAAGAGATCCTAAATCTTGGTCGGAACCTGGAAAGCTTTAAACCCAGAAAGG 1240
Db 1243 TGGGCGCATAGCAAGGAGCCCTGAGAAATTTGGGAAGAGCTTGAGAAATTTTCCCGAAAGG 1302
Qy 1241 TTAAAGATTTGCAATTTGATTTATAAAGGGA---CGACATTTGAACTGGTACCAATTTGGT 1297
Db 1303 TTCTCTGAGAGTTCCGATCGAGTTAAAGGGAATGATGAGTTTAAAGGTGATCCCGTTTGGT 1362
Qy 1298 CGAGAAAAAGATATGTCCTGGCAATTAATCTTACGATATACCAATTTGGAGTATGTCATT 1357
Db 1363 TCTGGAAGGAGAAATGTCTCTCGAAGCAGCATGGAATTTATGAATTTAGCTTTCTCTTG 1422
Qy 1358 ATAAATCTATTATATCATTTTAAATTTGGGA 1386
Db 1423 CTATCTCATTCACAGCTTTGATTGGGAA 1451

RESULT 13
US-10-953-349-32482
; Sequence 32482, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32482
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-32482

Query Match 8.6%; Score 148.4; DB 6; Length 1632;
Best Local Similarity 47.3%; Pred. No. 3.2e-30;
Matches 640; Conservative 0; Mismatches 681; Indels 33; Gaps 5;
Qy 113 CCTCAGGCGCCATGGAAGTTTCTCATCATPAGTAATCTTCTCATTTATTAATCTACTTCT 172
Db 137 CCGCCAGGCGCATGGCGCTCCCGGTCTATCGGTCTACCTGCACAT-----CTCGCGCGC 190
Qy 173 GATCTAGGCCATGAAGTTTATAGCCTTGGCTCAAATTTATGAGACCTGTTATGATCTT 232
Db 191 GCCCTTCCGCACCGTCTCTGCGGACCTCGCGCGCGCCACGCGCCGCTCATGATGCTC 250
Qy 233 CAAATTTGGCAAGTTTTCAGCTGTGTCTTCTTTCAGCTGGAAGCAGCAAGAGGTTATG 292
Db 251 CGCCTCGGCGAGCTCGACGCGGTGTGGCTGTCCCTTGAGCGCGCGCGGAGATCATG 310
Qy 293 AAAAATCAGGCTGATGCCCTTCGCCCAAAGCCCTATGCTCTTGGACGCAAGATTTGTTT 352
Db 311 AAGACCCACGACGCTCTCTTCGCTCCAGGCTCTGACGTCGATGACGAGATGCGGTAC 370
Qy 353 TATATCGGAAGATGCTTGTGTTGCTTATGAGATCACTGGAGGCGAGATGAAGAA 412
Db 371 GCGCAGCGCGAGGGGCTCATCTTCGCGCCCTACGCGCAGCATGGCGGAGCTTCGCAAG 430
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAAAGTTTCAATCCTCCAGGTTTAATCCGAGAG 472
Db 431 ATATGACTGTGAGATCTTCAGCTCCCGCGCGTCCAGTCTTCGCCCCCGCGCGAG 490
Qy 473 GAAGAAATGGAGATGCCATCAATTCCTCCGTTTGGAAAGCGGATCT-----CGGTTC 526
Db 491 GAGGAGCTCGGTCACTCTCCGCTCCGTCGCGCGCGGCTTCTGCTCTGCTTTCGCGGTG 550
Qy 527 ATATTACAAAGATCATTTATGGCATATAATTTTCGATCATGATGAACATCCGTTGTT 586

Db 551 AACCTGAGCGAGCGCATATCGCGCTACGTCGGGATTTCTACGGTGGCGGCCATCGTCGCG 610
Qy 587 A---ATTGTAAGCAAAAGAAAGATTGCTGAGTGTGTCGGATGTCAGTCAATGAGCAGCG 643
Db 611 AGCCGGTTCRAAGCAGAGGGACACGCTACCTGAGATGCTGCAAGAGGACTCAAAATCGTG 670
Qy 644 ACAGATTTTGGCAGCGGACAGCGCTTTCCGAGCTGGAAATTTACTTCACTATATCATTTGA 703
Db 671 CCGGGATGACCTTCCCGCATTTTCCCTCTCGGCTCTCGTGGGCTCTCTTACGAGC 730
Qy 704 GCTGAGTCAAAACCCAGCGTTTGCATCAGGAGATTGACGATATATCTTTGAAGAGATTCTT 763
Db 731 GTCCCCGCGAGGATACAGCGCATAGCCAGACATGAAGCTGTTTCATGGACACCATCATC 790
Qy 764 AATGAACACAAAGCCAAATAAGCCTTTTG------AAGCGGATAACTTAATG 808
Db 791 CAAGAGCACAGGTGAACAGCGGCCCGGACTGCGAGCGCGACAAAGAAAGAGACTTGTCTC 850
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACCGTTCAGTGCAGTGCACAAACGAAAGC 868
Db 851 GACGTGCTCTGAGACTCCAAAGGAAGGGGACTCGCAGTATCCACTCACCACTGACAAAC 910
Qy 869 ATCAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
Db 911 ATCAAGCCGTCACTGACATGTTTGGCGCGGACGAGACGTCGCGGACGAGCTG 970
Qy 929 GAATGGTAAATGGCAGAGCTGATGAAATTCAACTGAATTAAGAAAAAGCACAAGANA 988
Db 971 CAGTGGCGGATGGCGGAGCTGATACGGAACCCGGGTTATGCGGAAGGCGCAAGACGAG 1030
Qy 989 GTTAGACAAGTATTGTTGGAATGGAAGATTTGATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1031 GTCCGGCAGCAACTCGCGCGGACGACAAAGGTGAGGAGGCGGACCTGGCAGATCTGCGA 1090
Qy 1049 TTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTTACATCTCTCCG---TTGCTTTGATT 1105
Db 1091 TACCTTGGTTTGTTCATCAGGAGACGCTGAGGATGCACCCCGCGCGCCATTGCTGTG 1150
Qy 1106 CCGAGGAGGTAGAGAAACAAACGAATTTGATGATATGAATTTATCCCGAACACTCGA 1165
Db 1151 CCGGTAGTTCGGGAGCCGCTGTGAGTTTCTGGGCTCGGAGCTCGCGAGGGGTCATG 1210
Qy 1166 ATTGTTGTAATGCTTGGCGGATAGGAGAGATCTTAATCTTGTTCGAACTCGGAAAG 1225
Db 1211 GTGATCGTGAACGCTGGCGGATCGGCATGACCCGCGCATCTGGAGGACACCCGAGGAG 1270
Qy 1226 TTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGACGCAATTTGAATG 1285
Db 1271 TTGACCCGAGCGGTTTCGAGCAGAAACGGAGGAGTTCAGGGGCGGAGCTTCGAGTTC 1330
Qy 1286 GTACCATTTGGTCAGGAAAGAAATATGTCCTGCGNTTACTTCACTATTACCAATTTG 1345
Db 1331 GTGCCGTTTCGGCGCGGAGGAGGATATGCCCGCATGGCTTTCGGGCTGGCGCACGCTG 1390
Qy 1346 GAGTATGTCATTATAAATCTATTATCATTTTAATTCGGAACCTGGCCGATGGAATTACA 1405
Db 1391 GAGTCGCGCTCGCGCGCGCTGCTGTTCCACTTCGACTGGAGCTGCCAGGCGGGCGGCG 1450
Qy 1406 CCTCAAAACACTTGATATGACTGAAGCTATTGGCG 1439
Db 1451 GCCGAGGATCTGGACATGACCGAGGAGTTTGGCG 1484

RESULT 14
US-10-953-349-32468
; Sequence 32468, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32468
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-32468

Query Match 8.6%; Score 148.2; DB 6; Length 1766;
Best Local Similarity 47.5%; Pred. No. 3.7e-30;
Matches 614; Conservative 0; Mismatches 658; Indels 21; Gaps 5;

QY 114 CTCAGGGCCATGGAAGTTTCTTATCATAGGTAATCTTCTCATTTATTTACTCACATTCG 173
Db 183 CGCTGGCGCGTGGCAGCTGCGGTGATCGGCAGCTGCACCACTCGTCTGGCGGGC 242
QY 174 ATCTAGGCCATGAACGTTTATAGACCTTGGCTCAAAATTTATGACCTGTATAGTCTTC 233
Db 243 AGCTCCCTCACCGCGCGATGCGGACCTGGCGGACGCTACGGGCGCGGCTGCTCTCC 302
QY 234 ABAATGGCCAACTTTCAGCTGTGTCTATTTCTTACCTGACCTGAACAGCCAAAGAGTTATGA 293
Db 303 AGCTGGCCAGGTGAAGACGCTGGTGTCTCTCGGAGGGGGCGCGGAGGTGATGA 362
QY 294 AAACCTCAGGCTGATGCTTTCGCCCAACGCCCTATCTGTGACGACACAGATTGTGTTTT 353
Db 363 AGAACCCACGACCATGTTTGGCCACGCGCGCTGAGCACCACTGCGCGTCTCTCT 422
QY 354 ATAATCGGAAGATGTTGTTTGTTCATATGAGATCACTGGAGGCGAGATGAAGAAA 413
Db 423 ACGCGGGCCAGGACATCGTCTTCGCGCCCTACGGAGATACCTGGCGCCAGCTCCGCAAGA 482
QY 414 TTTGGATACTTGAATTTCTGAGTGCCAAAAGATTCAATCTCCAGTTAATCCGAGAGG 473
Db 483 TCGCGGTGCCAGCTCTTACCGCGCGCGGTCTCTCTTTCGCGGCATCCGCGAGG 542
QY 474 AAGAAATGGAGGATGCCATCACATTCCTCCG-----TTCGAAAGCCGGATCTCCGGTCA 527
Db 543 AGGAGTCCCCACGGCGCTCCGCGTCTCGCGAGGCTCGGCGGCGCGCGCCCGTGG 602
QY 528 ATATTACAAAGATCATTTATGGCATTAATAATTTGATCATGATAGAAATCCGTTGGTA 587
Db 603 AGATGCGCGGTGTCTATCCAGCTCGTACGCGACAGCAGCGCGCGCGCTCATAGGCG 662
QY 588 A---TTGTAAGCAAAAGAAAGATTGCTGAGTGTTCGCCGATGCAGTCAATGAGGACGA 644
Db 663 ACCGGTGCAGGAGCGCGACGCGTTCCTCCGGAGCTCGACCGCATCGTGCAGCTCGGT 722
QY 645 CGAGTTTTGGCACCGCAGACGCTTTTCCGACGTGGAAATTACTTCACTATATCATTTGGAG 704
Db 723 CGGGGTCAACCTGGCCGACATGTGGCGCTCGTCGACGCTTGGCGGTGGCTCAGTGGAG 782
QY 705 CTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATACCTTGAAGAGATTCTTA 764
Db 783 CCGAGGAGTGGCCACACCTGTACACCATGCTCGACGGCATCGTTGAGGAGCACCT-- 840
QY 765 ATGACACAAAGCCCAATAGCCTTTTGAAGCGGATAAATTAAATGATGTTCTATTGAATC 824
Db 841 -GGAGAGGACGACGCGCGGAGGCGCCACGCGGAGACTTGTCTGACGCTGCTCTCAAGA 899
QY 825 TTCAAAATAAATGAAAACGTTCCAGTGGCCAGTGACAAACGAAAGCATCAAGCATCCGTTT 884
Db 900 TCCAGAGGAGGGTGGTCTCAAGTTTCCCATCCACATGACGCGCGTCAAGGCATCATCT 959
QY 885 TCGAAATGTTTATCGCGGAGCGGAAACAACTTCGAAAGCTACAGAAATGGGTAATGGCAG 944
Db 960 TGGAGGTATTTCTCTGCGGCGAGTGAACAAACAACTACGACAATTTGAGTGGGCGCATATCAG 1019
QY 945 AGCTGATGAAAATCCAACCTGAACCTAAGAAAGGACAGAAAGATTAGACAGTATTGG 1004
Db 1020 AGCTGATCAACAAACCGGATGGCCATGTCAGAAAGGCGACAGATGAGGTGCGACAGCCTCC 1079

QY 1005 GTCAAAATGGGAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTGG 1064
Db 1080 ATGCTAGTGAACCGGTGTCGAGCAGCCCTAAGCAGAGCTCCCATACCTCGCTTGGTCA 1139
QY 1065 TTAAGAAAACCTTAAGATTACATCTCCGTTGTCTTGAATTCGAGGAGGTGAGAGAAA 1124
Db 1140 TCCGAGAGACGCTGCGGTACACCCGCCCTGCGGTGTGTGTTCCGCGAGTGCAGGAAC 1199
QY 1125 CAACAGCAATTGATGGATATGAATTCATCCGAACTCGAATTTGTTGAATGCTTGGG 1184
Db 1200 CGTGCCAAAGTGCAGGATACGAGCTGACGCGGGGACGAGGTGTTGGTCAATGCTTGGG 1259
QY 1185 CGATAGGAAGAGATCCTAATACTTGTGTCGA---ACCTGGAAGATTTAACCCAGAAAGGT 1241
Db 1260 CGTAGGCCCGCAGCAGCGCTATTGGCCCGCAGCGCTCGAGAGTTCCGGCCGGAACGGT 1319
QY 1242 TTAAGATTGTGA-----ATTGATTAAGAGGACGACATTTGAACCTGGTACCATTTG 1295
Db 1320 TCGAGGAAGAAGCAGCAAGGCGAGACTTTGGGGGTGGTGACTTTGGGTTCTTGGCGTTTG 1379
QY 1296 GTGCAGGAAAAGAAATATGCTCTGGCATTTACTTCAGCTATTACCAATTTGGAGTATGCA 1355
Db 1380 GCGCGCGCGCAGGATGTGCCCTGGGATGGCGTTTGGCTCGCCGCGTCGAGCTCCCGC 1439
QY 1356 TTATAAATCTATTATATATCAATTTAATTGGGAAC 1388
Db 1440 TTGCAAGCATGCTCTTCCACTTTGACTGGAAGC 1472

RESULT 15

US-10-953-349-12509
; Sequence 12509, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12509
; LENGTH: 1727
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (562)..(563)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1481)..(1481)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1484)..(1484)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1506)..(1506)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-12509

Query Match 8.5%; Score 147.2; DB 6; Length 1727;

Best Local Similarity 48.0%; Pred. No. 6.7e-30;

Matches 657; Conservative 0; Mismatches 675; Indels 36; Gaps 7;

QY 97 GAAGAAACAGATCCACCTCCAGGCGCCATGGAAGTTTCTATCATAGGTAATCTTCTCA 156
Db 122 GAATCAAGAAATTAACCACTGGTCCAAAGGGTTGCCAATTTTGGGAAGCCTTCACAA 181
QY 157 TTTATTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTTGGCTCAAATTTATGG 216

	Db	182	GTTGGGACCAAAATCCTCATCGTGACC-----TGCACAAAATAGCCCAAATAATGG	232
	Qy	217	ACCTGTATTAGAGTCCTCAAATTGGCCAAAGTTTCAGCTGTGTGTCATTTCTTTCAGCTGAAGC	276
	Db	233	ACCTGTCAATGCACTTAGCTTTAGGTTTTGTGCCGACCAFATAGTTGTGTTCTTACCCAAATC	292
	Qy	277	AGCCAAAGAGGGTTATGAAAACTCAGGCTGATGCCCTCGCCCAACGCCCTATCGCTTTGGGA	336
	Db	293	TGCTGAATTGTTCTCCTCAAGACCCATGACCTTGCTTTGCTAGTAGACCACTGTTCTGTGGC	352
	Qy	337	CGCACAGATTGTGTTTTATTAATCGGAAGATGCTCTGTTTGTCTTCATATCGAGATCACTG	396
	Db	353	GGATCAATATACATCTCTTTGGGGCGCAGAGAACTTAGGCTTTGCTGGAATATGTTCTTATTG	412
	Qy	397	GAGCGAGATGAAGAAAAATTTGGATACTTTGAATTTCTGAGTGCCAAAAAAGTTTCAATCCTC	456
	Db	413	GCGCAACATGCGCAAGATGTGCACATTTGGAATTTGCTAAGCCAATCCAAAATTAATCTCCTT	472
	Qy	457	CAGGTTAATCCGAGAGGAAGAAATGAGAGATGCCATCACATTCCTCGTTTCGAAAGCC--	514
	Db	473	CAGACGCATGAGGGAAGAGGAGCTTGACCTTTTGATCAAGCTTGTAAGAGAGGCGACCAA	532
	Qy	515	----GGATCTCCGGTCAATATTACAAGATCATTTATGGCATTTATAATTTTCGATCATGAT	570
	Db	533	TGATGGAGCTGCTCTTGATCTCAGTGTCAANNITGCAACACTCATTTGCAGACATGCTWTG	592
	Qy	571	AAGAACATCCGTTGG---TAATTGTAGCAAAAGAAAGATGCTGAGTGTTGCCGATGC	627
	Db	593	TAGAATGATTTTAGGNAAGAAGTACATGGAACAGGACATGTGTGGGAGAGGGTTCAAGGC	652
	Qy	628	AGTCAATGAGSCACGACGAGTTTGTGCACC-----GCAGACGCTTTTCCGACGTTGA	680
	Db	653	TGTGATTTCAAGAGCAATGCGTTTATTAGCAATCCTCTAACTAGGAGATTACATTCCTTA	712
	Qy	681	AAATTACTTCACTATATCATTTGGAGCTCAGTCAAAACC--CAGGCGTTTGCATCAGGAGAT	738
	Db	713	CATTGGTGCAATTACCTTCAAGGCCCTAACGAAGCGCTTCAAAGTACTTTACGAATCTT	772
	Qy	739	TGAGATATACTTGAAGAGATTCTTAATGAACACAAAGCCAAATAAGCCTTTTGAAGCGGA	798
	Db	773	CGATGACTCTTCTCAGAAAGATTATGTATGATGACCATGGAATCGAGAAAGGGAGAAACAA	832
	Qy	799	TA-----ACTTAATGGATGTTCTATTGAATCTTCAAAAAATGGAAACGTTCCAGTGCC	852
	Db	833	GACCAAGGATTTTGTGNATGTCATGTTGGGCTTTTATAGTACTGAAGAACTGNAATACCG	892
	Qy	853	AGTGACAAACGAAAGCATCAAAAGCATCCGTTTTCGAAATGTTTACTCCGGGAGCGAAAC	912
	Db	893	CATTGAACGGTGCCAATATCAAAAGCTATATTGCTGGATAGTTGGCTGTTTCAATGGATAC	952
	Qy	913	AAC TTCGAAGACTACAGAAATGGGTAAATGGCAGACTGATGAAAATCCAACTGNACTAAG	972
	Db	953	TTTCACTACAGCAATTGAGTGGACACTTTCAGAGCTACTAAAAATCCAAAGGGTGATGAA	1012
	Qy	973	AAAAGCAACAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAAGTTGATGAATCAAG	1032
	Db	1013	GAAACTCCAAATGGAAATGGAACTGTGTGGGTATGAAGGGAAGGTGGGGHAATCCGA	1072
	Qy	1033	ATTTTCATGATTTGAAATTTCTTCAAAGTTAGTGGTTTAAAGAACTCTAAGATTTACATCTCC	1092
	Db	1073	CCTGGACAAGTTGAAGTATTTTGAGATGTTGTGTAAAGGAAAGCATGAGGCTCCATCCAGT	1132
	Qy	1093	GGTTGTCWTG---ATTCCGAGGAGTGTAGAGAAACAAACGAATTTGATGGATATGAAT	1149
	Db	1133	GGTACCAATTTGCTAATACCAACCAATCCACGGAAGATTTGCATAGTCGGAGATTTTTTCAT	1192
	Qy	1150	TCATCCGAACTCGAATGTTGTGTAATGCTTTGGCGATAGGAAGAGATCCTAATACTTG	1209
	Db	1193	ACCCAAAAAATCAAGGGTCAATAATAATGCAATGGGCAATTAATGAGAGCCCAAGTGCATG	1252
	Qy	1210	GTCGGAACCTCGAAAGTTTAAACCCAGAAAGGTTTTAAAGATTGTGCAATTTGATTATAAAGG	1269

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 09:04:09 ; Search time 351 Seconds
(without alignments)
9238.267 Million cell updates/sec

Title: US-10-759-813-1

Perfect score: 1733

Sequence: 1 gcaataaaggaaatggagc.....tcttttcaatccgaaaaa 1733

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfilesi.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	399.4	23.0	1781	US-09-499-302A-1
2	369.8	21.3	1665	US-08-313-075A-29
3	367.4	21.2	1666	US-07-912-900-24
4	367.4	21.2	1666	US-08-285-309-24
5	367.4	21.2	1666	US-08-502-046-24
6	308	17.8	1691	US-08-948-564-3
7	253.2	14.6	1515	US-09-292-768-5
8	245.4	14.2	1762	US-08-881-784-5
9	245.4	14.2	1762	US-09-292-768-1
10	245.4	14.2	1762	US-09-292-768-65
11	243.8	14.1	1665	US-09-292-768-67
12	243.8	14.1	1762	US-09-292-768-63
13	242.2	14.0	1665	US-08-881-784-8
14	242.2	14.0	1665	US-09-292-768-3
15	242.2	14.0	1665	US-09-292-768-69
16	234.2	13.5	1762	US-09-172-339-5
17	224.2	12.9	1838	US-08-948-564-1
18	169	9.8	1488	US-09-689-783A-1
19	164.2	9.5	1812	US-08-313-075A-37
20	164.2	9.5	1824	US-08-606-505B-1
21	164.2	9.5	1824	US-09-616-990-1
22	162.6	9.4	1657	US-08-948-564-11
23	162.4	9.4	1755	US-07-912-900-29

24	162.4	9.4	1755	2	US-08-285-309-29	Sequence 29, Appl
25	162.4	9.4	1755	2	US-08-502-046-29	Sequence 29, Appl
26	161.4	9.3	1757	2	US-08-313-075A-49	Sequence 49, Appl
27	161.4	9.3	1757	3	US-09-142-108C-26	Sequence 26, Appl
28	159.6	9.2	1789	3	US-09-142-108C-1	Sequence 1, Appl
29	151.4	8.7	1539	3	US-10-142-231-53	Sequence 53, Appl
30	151.4	8.7	1539	5	US-10-884-115-53	Sequence 53, Appl
31	148.4	8.6	1722	3	US-09-033-055A-1	Sequence 1, Appl
32	146.6	8.5	1812	2	US-07-912-900-28	Sequence 28, Appl
33	146.6	8.5	1812	2	US-08-285-309-28	Sequence 28, Appl
34	146.6	8.5	1812	2	US-08-502-046-28	Sequence 28, Appl
35	142	8.2	1545	3	US-10-142-231-55	Sequence 55, Appl
36	142	8.2	1545	5	US-10-884-115-55	Sequence 55, Appl
37	139.6	8.1	1764	3	US-09-947-027-3	Sequence 3, Appl
38	137	7.9	1634	3	US-09-126-420A-2	Sequence 2, Appl
39	137	7.9	1737	3	US-09-126-420A-1	Sequence 1, Appl
40	133.2	7.7	1724	3	US-09-672-785-1	Sequence 1, Appl
41	127.2	7.3	1835	3	US-09-564-808-5	Sequence 5, Appl
42	127.2	7.3	1884	3	US-09-564-808-3	Sequence 3, Appl
43	126.4	7.3	1667	3	US-09-142-108C-22	Sequence 22, Appl
44	125.6	7.2	1838	2	US-09-091-432-1	Sequence 1, Appl
45	125.6	7.2	1838	3	US-09-387-663-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-499-302A-1
; Sequence 1, Application US/09499302A
; Patent No. 6369212
; GENERAL INFORMATION:
; APPLICANT: BOUNG-JUN, OH
; APPLICANT: MOON, KYUNG KO
; APPLICANT: YOUNG, SOON KIM
; TITLE OF INVENTION: A CYTOCHROME P450 GENE HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 10324/P64443USO
; CURRENT APPLICATION NUMBER: US/09/499,302A
; CURRENT FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1781
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-499-302A-1

Query Match	23.0%	Score 399.4;	DB 3;	Length 1781;
Best Local Similarity	56.3%	Pred. No. 8.6e-98;		
Matches	826;	Conservative	0;	Mismatches 611; Indels 30; Gaps 3;
QY	99	AGAAACAGAAATCCACCTCCAGGGCCATGGAAGTTTCTTATCATAGTAAATCTTCTCATTT	158	
Db	86	AAAGCTAAACTTACCTCTCTGGTCCATGGAATACCTTTTATTGGAAGCCATACACT	145	
QY	159	TATTACTCATTCTCTGATCTAGGCCATGAACGTTTTAGAGCCCTGGCTCAAAATTTATGGAC	218	
Db	146	TGGCAGTGGCAGGTCCACTTCTCATATGGCCCTAAAAAATCTAGCAAAATTTATGGGC	205	
QY	219	CTGTTATGAGTCTTCAAAATTTGGCCAAAGTTTTCAGCTGTTGTCTATTCTTTCAGCTGAAGCAG	278	
Db	206	CGCTCATGCACTTACGATCTGGGGAAATTCCTACGTCATCATTTTGGTCCCGCAATGG	265	
QY	279	CAAAAGAGGTTATGAAAACTCAGGCTGATGCTTCGCCCAAGCCCTATFCGTCCTTGGAGC	338	
Db	266	CGAAGGAAGTACTAAAAAATCAGCAGCTCGCTTTCGCAACAGGGCGGAAACTTGTGGTG	325	
QY	339	CACGATTTGTTTATTAATCGAAAGATGCTGTTGCTTCATATGAGATCACTGA	398	
Db	326	CTGACATCGTCCATTTATGATAGTACGATATAGCATTTTCTCCATATGTTGAATACTGA	385	
QY	399	GSCAGATGAAGAAAATTTGGATCTTGAATTTCTGAGTGCCCAAAAAGTTCAATCTCTCA	458	

386	Db	GGCAGATTCTGTAATAATTTGGATACTCTGAACTCTCTTAGTGCAGAGATGGTCAAAATCTTTTA	445
459	Qy	GGTTAATCCGAGAGGAATAATGGAGATGCCATCACATTCCTCCGTTTGGAAAGCCGGAT	518
446	Db	GCTCAATTCGCCAGGATGAGCTGCGATGATGGTCTCATCTATACGAAACCATGCCAAAT	505
519	Qy	CTCCGGTCAATATTTACAAGATCATTTATGGCAATTAATTTTCGATCATGATGAAGACAT	578
506	Db	TTCCGTCACACTTACAGACAAAATAATTTTGGTTTCAAGTTCGGTAACTTGTAGATCAG	565
579	Qy	CCGTTGGTAA---TTGTAAGCAAAAGAAAGATTGCTGAGTGTTCGGATGCGAGTCAATG	635
566	Db	CTCTGGGAAAAATATGTCGTGACCAAGACAAACTGATAAATTTTCATGAGGGAATAATAT	625
636	Qy	AGGCAGCGACGAGTTTTGGCACCGCACAGCGTTTTTCGACGTGGAAATTACTTCACTATA	695
626	Db	CATTGACAGGTGGATTTAGTATTGCTGATTTTTCCTACATGGAATAATGCTACATGATG	685
696	Qy	TCATGAGCTGAGTCAAAACCCAGGCGTTTTGCATCAGGAGATTGACGATATACTTTGAAG	755
686	Db	TTGGTGGTTCAAAAACCTAGACTGCTGAAGGCTCATCTGTAATAATCGATGAGATTTTGAAC	745
756	Qy	AGATTCCTTAATGAACACAAAGCCAAATAAGCT-----TTTTG	791
746	Db	ATGTTAGTGAATGACCAAAACAGNATCGAGGAGTGCCTAAAGGTAATGGCGAATTTG	805
792	Qy	AAGCGGATAACTTAATCGATGTTCTATTGAACTTTCAAAAAAATGGAAACGTTTCCAGTGC	851
806	Db	CGCGTGAAGATTTCATCGATGTTTGTCTAAGGGTTCGAGAAGTGGAGAAAGTTTCAAAAT	865
852	Qy	CAGTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTCGCGGGGCGGAAA	911
866	Db	CCATCAGCGGATGACAAATATCAAAATCAATATTAGTGGACATGTTCTCCGCTGGATCTGAAA	925
912	Qy	CAACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAAC	971
926	Db	CGTCATCGACAACTATAATTTGGSCATTAGCTGAAATGATGAAGAACCAAGTGTCTTAG	985
972	Qy	GAAAGCAACAAGAAAGATTGACAAAGTATTGGTGAATGGGNAAGTTGATGAATCAA	1031
986	Db	CAAAAGGCACAAGCTGNAAGTGAGACAGCTTTGAAGGAAAAAGAAAGTTTTCACAAATG	1045
1032	Qy	GATTTTCATGATTGCAAAATTCCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTAACATCCTC	1091
1046	Db	ATCTTGATGAGTTGAGTACTTTGAAGTTAGTAACTCAAGAAACCTTAAGGATGCACCTC	1105
1092	Qy	CGGT---TGTCTTGAATCCGAGGAGTGTAGAGAAAACAACGAAATTTGATGGATATGAAA	1148
1106	Db	CAATTCCTCTATTAGTCCTAGAGAAATGATGAAGGATACAAAGATTTGACGGGTACAA	1165
1149	Qy	TTCAATCCGAACACTCGAATGTTCTGATGCTTCGGCGATAGGAAGAGATCCCTAATACTT	1208
1166	Db	TACCTTTCAAACTCGAGTCAATGTTAATGCAATGGCAATGGACGAGATCCTGAAAGTT	1225
1209	Qy	GGTCGGAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTATAAG	1268
1226	Db	GGGATGACCTGMAAGCTTTTCCCGACAGAGATTTCGAGATAGTTCCTGTTGACTTCTCTG	1285
1269	Qy	GGACGACATTTGAACTGGTACATTTGGTGACGAAAAAGAAATATGCTCTGGCAATTA	1328
1286	Db	GAAGCCATCATCAATTTATTCATTTGGTGGGAAAGAGATTTGCTCGGAATGCTTT	1345
1329	Qy	CAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTAAATTTGGGAAC	1388
1346	Db	TTGGTTAGCCAAATGTTGGACAAACCATTAGCTCAATTACTTTATCACTTCGATCGGAAC	1405
1389	Qy	TGGCGGATGGAATTAACCTCAACACATTGATATGACTGAAGCTATTGGCGGCTCTCA	1448
1406	Db	TCCCTAATGGACAAAGTCAAGAAAAATTTGGACATGACGGAGTCACTCGGAATTTCTGCA	1465
1449	Qy	GGAAAAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTCAATA	1508

Db 1466 CAGAAAGGATGATCTTGTTTGAATGCCACCCTTATGATCTTGAATGATGACACA 1552
 Qy 1509 TTTCTTGAATACATAGGAGGGTTGAAA 1535
 Db 1526 GTTGTAGAAATAAAAAAGAGGGAGAAA 1552

RESULT 2
 US-08-313-075A-29
 ; Sequence 29, Application US/08313075A
 ; Patent No. 5639870
 ; GENERAL INFORMATION:
 ; APPLICANT: Holcon, Timothy A.
 ; APPLICANT: Cornish, Edwin C.
 ; APPLICANT: Tanaka, Yoshikazu
 ; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
 ; TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 58
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/313,075A
 ; FILING DATE: 30-NOV-1994
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: AU PL 1538/92
 ; FILING DATE: 27-MAR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: AU PL 6698/93
 ; FILING DATE: 07-JAN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: AU PCT/AU93/00127
 ; FILING DATE: 25-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DiGiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 9433
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1665 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3..1432
 US-08-313-075A-29

Db 138 CATATGGTA-----GGTGGACTTCCACACCATGTCCTTTAGAGATTTAGCCAAAAATAT 191
QY 215 GGACCTGTTATGAGTCTTCAAAATGGCCCAAGTTTCAGCTGTGTCATTTCTTCAGCTGAA 274
Db 192 GNCACAAATATGACACTTCAACTAGTAAATTTCTGCCGTGTAGTTACTTCTCCTGAG 251
QY 275 GCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTG 334
Db 252 ATGGCAAGAAAGTAGTAAAAACTCATGACCTTGCATTTGCATATAGGCCCTAACTCTTA 311
QY 335 GAGGCACAGATTTGTTTATATCGAAGATGCTTGTTCCTTCATATGAGATCAC 394
Db 312 GCAATGAGATTTGCTGCTATATAGTTCCAGCAATTCGCTTTTCCCGGTGATTAC 371
QY 395 TGGAGGCAGATGAAGAAAAATTTGGATCTTGAATTTCTGAGTGCACAAAAAGTTCAATCC 454
Db 372 TGGAGGCAATGCGTAAATTTGTGTTTGGAGTGCCTTAGTGCACAAAAATGTCCGGTCA 431
QY 455 TCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCGATCACAATTCCTCCGTTTCGAAGCC 514
Db 432 TTTAACTCGATTAGACGAGATGAATACTTCTTATGATCGATTTTTCGATCATCTTCT 491
QY 515 GATCTCCGCTCAATATACAAAGATCATTTATGCAATTAATTTTCGATCATGATAAGA 574
Db 492 GGTAAAGCCAGTTAATATAACAGAAAGGATCTTTTCATTCACAAGCTCTATGATTTGTGA 551
QY 575 ACATCCGTTGGTAAATG---TAAGCAAAAGAAAGATTCGCTGAGTGTCCCGATGCAGTC 631
Db 552 TCAGTATTTGGGAAAGAAATAAGGAAAGAGATGTATAGCATGTGAAAAAATG 611
QY 632 AATGAGGCAGCAGGATTTTGGCACCGCAGCGCTTTTCCGACGTGGAAATTTACTTTCAC 691
Db 612 ACAGGCTTAATAGATGGGTTCGATGTGCTGCATATTCCTTCGTTGAGGTTCTTTCAT 671
QY 692 TATATCATTTGAGCTGAGTCAAAACCCAGCGCTTTTCATCAGAGATTCACATATCTT 751
Db 672 GTACTAATCGGTATGAAGGGTAAAAATATGATGTTTCATCGTAAGGTAGATCTATTGTT 731
QY 752 GAAGAGATTTCTTAATGAACACAAAG-----CCAAATAGCCTTTT 790
Db 732 GAGGAGTCACTGANTGAGCACAAAGAACTCTTCGACTGGCAAGCCAAATGGTGAAGTG 791
QY 791 GAAGCGGATTAATGATGTTCTTATGAAATCTTCAAAAAATTTGAAACGTTCCAGTG 850
Db 792 GGAGGAGAAGATTTAATGATGTTATGCTTAAGACTTAAAGGAAGAGGAGACCTTCAACTT 851
QY 851 CCAGTGCACAAACGAAGCATCAAGCATCCGTTTTCGAAATGTTTTCGCCGGGAGCGAA 910
Db 852 CCAATCACAAAATGACACATCAAGGCCATTTTAAATGACATGTTTTCGCTCGGGGAACAGAA 911
QY 911 ACAACTTCGAAAGCTACAGATGGGTAAATGCGAGAGCTGATCAAAAAATCCAACTGAATTA 970
Db 912 ACTTCATCAACACAAATTAACCTGGGCCATGGTAGAACTGATGAAAAATCCAAAGTGTATC 971
QY 971 AGAAAGCACAAGAGAAGTTAGACAAGTATTTGGTGAATGGGAAAAAGTTGATGAATCA 1030
Db 972 GCGAAAGCTCAAGCAGAGAGTAAGAGAAGTCTTCAAAAGGGAAGAAACTTTTCGATGAAGAT 1031
QY 1031 AGATTTTCATGTTGAAATCTTCAAGTTAGTGGTTAAAGAACTCTTAAGATTACATCT 1090
Db 1032 GATATCGAGGAGCTGAATTTACCTTAAGTTAGTCAATTAGAGAACTTTAAGACTCCACCT 1091
QY 1091 CCGGTT---GTCTTCAATCCGAGGAGTGTAGAGAAACAAACGAAATTTGATGATGAA 1147
Db 1092 CCACTTCACATTTTCTTCCAGAGAAATGTTCGAGAGAAACAGAAATTAATGGCTACACT 1151
QY 1148 ATTCATCCGAACACTCGAATTTGTGAAATGCTTCGGCGATAGGAGAGATCCTTAATACT 1207
Db 1152 ATTCCTTTAAATACCAAGTCTAGTTAATGTTTGGCTTATTTGGAAGAGATCCAAATAT 1211
QY 1208 TGGTCGGAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGCGAATGATTATAAA 1267
Db 1212 TGGGATGATGCAAGAAAGCTTTAAGCCTGAGAGATTTGAACATACTCTTTGAAATTTGCT 1271

QY 1268 GGGACGACATTTGAATCGTGTACCATTTGGTGAGGAAAAAGAAATATGTCTCGCATTA 1327
Db 1272 GCAATAATTTTGAATATATCTTCTTTTGGTAGTGGAAAGAGGATTTGCCCGGAATATCA 1331
QY 1328 TCAGCTATTACCAATTTGGAGTATGTCAATTAATAATCTAATATATATCATTTTAATTCGGAA 1387
Db 1332 TTTGGTTTAGCTAATGTTTATCATCCATTTGGCTCAATTTGTATCATTTTCGATTGAGA 1391
QY 1388 CTGGCGGATGGAATTAACCTCAACACATTTGATATAGCTGAAGCTAATTCGCGGTGCTCTC 1447
Db 1392 CTTCCTACTGGGGTCGACCCAAATGACTTTTGAATTTGACT-AGTTAGCTGGAGTAACTACT 1450
QY 1448 AGCAAAAAAATAGATCTTAAGTTGATTCCTATTTCCATAT 1486
Db 1451 GGTAGAAAAAGAGACCTTTACTTGAATTTCACTCTCTTAT 1489

RESULT 3

US-07-912-900-24
; Sequence 24, Application US/07912900
; Patent No. 5349125
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwina C.
; APPLICANT: Kovacic, Filipa
; APPLICANT: Tanaka, Yoshikazu
; APPLICANT: Lester, Diane R.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/912,900
; FILING DATE: 19920713
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-912-900-24

Query Match 21.2%; Score 367.4; DB 2; Length 1666;
; Best Local Similarity 55.7%; Pred. No. 3.8e-89;
Matches 824; Conservative 0; Mismatches 621; Indels 34; Gaps 5;
QY 35 TTTCCGAGCATTTTAATAAGTTTTCGCTTGTGTTTAAATCTTAGTAGTAGTAGAGTTG 94
Db 19 TTTCTTCTTCTTATGTTATTTTCCCTCATTTTCAATTAAGAAAAATGAAGAAATCCAAT 78
QY 95 TGAAGAAACAGAAATCCACCTCCAGGGCCATGGAAGTTTCTCTATCATAGTAATCTTCT 154

Db 79 TGTCAAAACCAAAATTTGCTCCAGGCCCATGAAAGTACCTTTTCITGGAGCTTGCTT 138
Qy 155 CATTTTATCTACACTTCTGATCTAGCCATGAACGTTTTAGAGCCTTGCGTCAAAATTTAT 214
Db 139 CATATGGTA-----GGTGGACTTCCACCAATGTCCTTAGAGATTAGCCAAAATAT 192
Qy 215 GGACCTGTTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAA 274
Db 193 GGACCAATATGACACTTCAACTAGTAAATTTCTGCGGTGTGATTTACTTCTCCTGAG 252
Qy 275 GCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCTTCGCCCAAGCCCTATCGTCTTG 334
Db 253 ATGGCAAGAAAAGTAGTAAAACTCATGACCTTGCAATTTGCAATAGGCCCTAAACTTTCTA 312
Qy 335 GAGGCACAGATTGCTTTTATAATCGGAAAGATGCTTGTGTTTCTCATATGGAGATCAC 394
Db 313 GGCATTGAGATTGCTGCTATATAATAGTTTCAGACATTTGCCCTTTTCCCGTATGGTATAC 372
Qy 395 TGGAGGCAGATGAAGAAAATTTGGGATACCTTGAATTTCTGAGTGCACAAAAGTTTCAATCC 454
Db 373 TGGAGGCAATCGTAAATTTGTGATTTGGAAGTCTTAGTGCCAAAATAGTCCGGTCA 432
Qy 455 TCCAGGTTAATCCGAGAGAGAAATGGAGGATGCCATCACATTCCTCGTTCGAAAGCC 514
Db 433 TTTAACTCGATTAGACGAGATGAAATACCTTCTTATGATCGATTTTTCGCGATCATCTCTC 492
Qy 515 GGAATCCCGTCAATATTAACAAGATCAATTTATGCAATATATAATTTTCGATCATGATAAGA 574
Db 493 GGTAAAGCCAGTTAATATAACAAGAGGATCTTTTCAATCAAGCTCTATGATTTGTAGA 552
Qy 575 ACATCCGTTGGTAAATG---TAAGCAAAAAGAAAGATTGCTGAGTGTGTGCCGATCAGTC 631
Db 553 TCAGTATTTGGGAAAGRATAAGGAGAAAGCAATGATATACGACATGTGAAAAAATG 612
Qy 632 AATGAGCGACGACGAGTTTGGCCCGCAGACGCTTTTCCGACGTGGAAATTTACTTCAC 691
Db 613 ACAGGCTTAATAGATGGGTTCGATGTGGCTGCATATTCCTTCGTTGAGGTTTCTTCAT 672
Qy 692 TATATCATTTGAGCTGAGTCAAAACCCAGCGCTTTGCAATCAGGAGATTGACGATATACCT 751
Db 673 GTACTAATCGGTATGAGGGTAAATATGATGTTTCATCGTAAGGTAGATCTATGTT 732
Qy 752 GAAGAGATTCTTAATGAACACAAAG-----CCAATAAGCCTTTT 790
Db 733 GAGGAAGCTATGAATGAGCACAAAGAAACTCTTCGAACTGGCAAGACCAATGGTGAAGTG 792
Qy 791 GAAGCGGATACTTAATGATGTTCTATTGAACTTCAAAAATGGAACCGTTCCAGTG 850
Db 793 GGAGGAGAAAGATTAAATGATGTTATGCTAAGACTTAAGGAAGAGGAGACCTTCAACT 852
Qy 851 CCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTCCCGGAGCGAA 910
Db 853 CCATCACAATATGACAACACTAAAGCCATTTTATGACATGTTGCTCGGGAACAGAA 912
Qy 911 ACAACTTCGAAAGCTACAGAATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAAC 970
Db 913 ACTTCATCAACACAAATTAACCTGGGCCATGGTAGAATGATGAAAAATCCAGTGATTC 972
Qy 971 AGAAAGCAAGAGAGGTTAGACAAGTATTGTTGTAATGGGAAAGTTGATGATATCA 1030
Db 973 GCGNAAGCTCAAGCAGAGGTAAAGAGAGTCTTCAAGGGGAAAGAACTTTTCGATGAAGAT 1032
Qy 1031 AGATTTTCATGATTGAAATCTTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACATCT 1090
Db 1033 GATATCGAGAGCTGAAATTAACCTTAAGTTAGTCAATTAAGAAACTTTAAGATCTACCT 1092
Qy 1091 CCGGTT---GTCTTGATTCCGAGGAGTGTAGAGAAACACAGAAATGATGATGAA 1147
Db 1093 CCACTTCCACTTTTGTCTTCAAGAGAAATGTCGAGAGAGAAACAGAAATAAATGGCTACACT 1152
Qy 1148 ATTCATCCGAACTCGAATTTGTTGAATGCTTGGCGGATAGAGAGATCTTAATCT 1207

Db 1153 ATTCCTTTAAATACCAAGTCATAGTTAATGTTTGGGCTATTGGAGAGATCCAAATAT 1212
Qy 1208 TGGTCGGAACCTGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAATA 1267
Db 1213 TGGGATGATGCAGAAAGCTTTAAGCCTGAGAGATTGAAACATAACTCTTTTGAATTTTGT 1272
Qy 1268 GGGAGCAGATTGAACTGGTACCAATTTGGTGCAGGAAAGAAATATGCTTGGCATTACT 1327
Db 1273 GGCAATAATTTGAATATCTTCTCTTTTGGTAGTGGAGAGGATTGCCCCGGAATATCA 1332
Qy 1328 TCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATATCATTTTAAATTGGGAA 1387
Db 1333 TTTGTTTAGCTAATGATGATCATCCATTGCTCAATTTGTTGATCAITTCGATTGGAGA 1392
Qy 1388 CTGCCCGATGGAATTTACACCTCAACACTTGATATGATGAGCTATTTGGCGGTCTCTC 1447
Db 1393 CTTCCTACTGGGTGCGACCCAAATGACTTTGAAATTGACT-AGTTAGCTGGAGTAACTACT 1451
Qy 1448 AGGAAAAAATAGATCTTAAGTTGATTCCTATTCCATAT 1486
Db 1452 GGTAGGAAAAGAGACCTTTACTTGTATTTCCTCTTAT 1490

RESULT 4

US-08-285-309-24
; Sequence 24, Application US/08285309
; Patent No. 5569832
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwin C.
; APPLICANT: Kovacic, Filipa
; APPLICANT: Tanaka, Yoshikazu
; APPLICANT: Lester, Diane R.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/285,309
FILING DATE: 03-AUG-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8633Z
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1666 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLSCALE TYPE: DNA (genomic)
US-08-285-309-24

Query Match 21.2%; Score 367.4; DB 2; Length 1666;
Best Local Similarity 55.7%; Pred. No. 3.8e-89;
Matches 824; Conservative 0; Mismatches 621; Indels 34; Gaps 5;

Qy	35	TTTTCCGAGCATTTTAAATAGTTTTTCTGCTGTTGTTTAAATCTTTAGTAGTAGTCATGAGGTG	94
Db	19		
Qy	95	TGGAAGAAAACAGAATCCACCTCCAGGGCCATGGAAGTTTCCATCATAGAGTAAATCTTCCT	154
Db	79		
Qy	155	CATTATTTACTCACTTCTGATCTAGGCCATGAAGCTTTTAGAGCCTTTGGCTCAAATTTAT	214
Db	139		
Qy	215	GGACCTGTTATAGTCTTCAAAATGGCCAGTTTCAGCTGTTGTCATTTCTTCAGCTGAA	274
Db	193		
Qy	275	GCAGCCAAAGAGGTATGAAGATCAGGCTGATGCCCTTCGCCCAAGCCCTATCGTCTTG	334
Db	253		
Qy	335	GAGCACAGATTTGTTTATTAATCGGAAAGATGTTGTTTGTCTTCATATGAGATCAC	394
Db	313		
Qy	395	TGGAGGCAGATGAAGAAAAATTTGGATCTCTGAATTTCTGAGTGCAGAAAAAGTTCAATCC	454
Db	373		
Qy	455	TCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTCCTCCGTTTCGAAAGCC	514
Db	433		
Qy	515	GGATCTCCGGTCAATATATACAAAGATCATTTATGGCATTAATTTTCGATCATGATAAGA	574
Db	493		
Qy	575	ACATCCCTTGGTAATG---TAAGCAAAAAAGAAAGATGTGTGAGTGTGCGATGCGAGTC	631
Db	553		
Qy	632	AATGAGCGCGACGAGTTTTTGGCACCGCAGACGCTTTTCCGAGCTGGAAATTACTTTCAC	691
Db	613		
Qy	692	TATATCATTTGGAGCTGAGTCAAAACCCAGGCGTTTTCGATCAGGAGATTTGACGATATACTT	751
Db	673		
Qy	752	GAAGAGATTTTAAATGAACACAAAG-----CCAAATAAGCCCTTTT	790
Db	733		
Qy	791	GAAGCGATTAATTAATGGATGTTCTATTGAACTTTCAAAAAAATGAAAGCTTCCAGTG	850
Db	793		
Qy	851	CCAGTGACAAAACGAAAGCATCAAAGCATCCGTTTTCGCAATGTTTACTGCGCGGAGCGAA	910
Db	853		
Qy	911	ACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACCTGA	970
Db	913		
Qy	971	AGAAAACACNAGAAAGTTTAGCAAGTATTTGGTGAATTTGGGAAAAGTTGATGATCA	1030
Db	973		
Qy	1031	AGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAACTCTTAAGATTTACATCCT	1090
Db	1033		
Qy	1091	CCGGTT---GTCTTTGATTCGAGGGAGGTAGAGAAAACAACAGAAATTTGATGGATGAA	1147

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-502-046-24

Query Match      21.2%; Score 367.4; DB 2; Length 1666;
Best Local Similarity 55.7%; Pred. No. 3.8e-89;
Matches 824; Conservative 0; Mismatches 621; Indels 34; Gaps 5;

QY 35 TTTCCGACGATTTTAATAGTTTCTGCTGTGTTTAAATCTTAGTAGTAGTCATGAGGTG 94
Db 19 TTTCTCTTCTCTTATGTTATTTTCCCTCATTTTCAATTAAGAAATGGAAGAAATCCAAT 78
QY 95 TGAAGAAACAGAAATCCACTCCAGGCCATGGAAGTTTCTTATCATAGGTAATCTTCT 154
Db 79 TGTCAAAACCAAAATATGCTCCAGGCCATGGAAGTACCTTTTCTTGGAGCTTGCTT 138
QY 155 CATTTATTACTCAGTTCTGATCTAGCTAGCCATGAACGTTTGTAGAGCTTGTAGAGATTTAGCCAAAATAT 214
Db 139 CATATGGTA-----GGTGGACTTCCACACCATGTCTTAGAGATTTAGCCAAAATAT 192
QY 215 GGACCTGTTATGAGTCTTCAAAATGGCCAAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAA 274
Db 193 GGACCAATATGACCTTCAACTAGTAAATTTCTGCGTGTAGTTACTTCTCTGAG 252
QY 275 GCAGCAAGAGGTTATGAAACTCAGGCTGATGCTTGGCCCAAGGCCCTATCGTCTTG 334
Db 253 ATGGCAAGAAAGTACTAAAACTCATGACCTTGCATTTGCATATAGGCTTAACTTCTA 312
QY 335 GAGCACAGATGTTGTTTATTAATCGGAAGATGCTTGTGTTGCTTATATAGGATCAC 394
Db 313 GGCATTGAGATTTGCTGCTATTAATAGTTTGCAGACATGCGCTTTCCCGTATGGTATAC 372
QY 395 TGGAGCGAGATGAAGAAATTTGGTACTTGAATTTCTGAGTGCACAAAGTTCAATCC 454
Db 373 TGGAGCAATCGTAAATTTGTTATTTGGAAGTCTTAGTGCACAAAGTCTCCGGTCA 432
QY 455 TCCAGGTTAATCCGAGAGAAATGGAAGATGCAATCACATTTCTCGGTTGGAAGCC 514
Db 433 TTTAACTCGATTAGACGAGATGAAATACCTTCTTATGATCGATTTTGTGCGATCATCTCTC 492
QY 515 GCATCTCGGTCATATTAACAAGATCAATTTATGCGATTATATTTTCGATCATGATAAGA 574
Db 493 GGTAGCCAGTTAATATAACAGAGGATCTTTTATTCACAGCTCTATGATTTGTAGA 552
QY 575 ACATCCGTTTGGTAATTTG---TAAGCAAAAGAAAGATTGCTGAGTGTGCGGATCGAGTC 631
Db 553 TCAGTATTTGGGAAAGAAATAAAGGAGAAAGCAATGTATACGACATGTGAAAAAATG 612
QY 632 AATAGGCGACGACGAGTTTGGCACCGCAGAGGCTTTCCGACGTGGAATTAATCTCAC 691
Db 613 ACAGGCTTAATAGATGGGTTGCTATGCTGCTGACATATTCCTTCGTTGAGGTTTCTTCAT 672
QY 692 TATATCATTTGAGCTGAGTCAAAACCCAGCGTTTTCATCAGAGATTACGATATACAT 751
Db 673 GTACTAATCGGTATGAAGGGTAAATATGATGTTTCAATCGTAAAGTAGATGCTATTGTT 732
QY 752 GAAGAGATTCTTAATGAACACAAAG-----CCAAATAGCGCTTTT 790
Db 733 GAGGAGCTATGAATGAGCACAAGAAATCTTTCGAACCTGGCAAGACCAATGGTGAAGTG 792
QY 791 GAAGCGGATACCTTAATGATGTTCTTATTTGAATCTTCAAAATAAGAAACGTTCCAGTG 850
Db 793 GGAGGAGAAAGATTAAATGATGTTATTTGCTAAGACTTAAAGAAAGAGGAGACCTTCAACT 852
QY 851 CCAGTCACAAACGAAAGCATCAAGCATCCGTTTGTGCAATGTTTATGCTCGGAGCGGAA 910
Db 853 CCATCACAATATGACAACTAAAGCCATTTTAAATGACATGTTTCTCGGGGAGCAGAA 912
QY 911 ACAACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCAACTGAACATA 970
Db 913 ACTTCATCAACAACAAATTAACCTGGGCGATGGTAGAACTGATGAAAAATCCAGGTGATTC 972
QY 971 AGAAAGCACAAGAAAGATTAGACAAGTATTTGGTGAATTTGGGAAATGGGAAAGTTGATGATCA 1030
```

```

Db 973 GCGAAGCTCAAGCAGAGGTAAAGAAAGTCTTCAAGGGAAAGAACTTTTCGATGAAGAT 1032
QY 1031 AGATTTTCATGATTTGAAATCTTCAAGTTAGTGTAAAGAAACCTCTTAAGATTAATCTCT 1090
Db 1033 GATATCGAGGAGCTGAATTTACCTTTAAGTTAGTCTATTAGAGAACTTTTAAGATCTCACCT 1092
QY 1091 CCGGTT---GTCCTTATTCGAGGAGTGTAGAGAAACCAACACGAAATTTGATGGATGAA 1147
Db 1093 CCACCTTCCACTTTTGTCTTCAAGAGAAATGTCGGAGAGAAACAGAAATAAATGGGCTACACT 1152
QY 1148 ATTCATCGAAGACTCGAATTTGTTGTAATCTTGGGCGATAGGAAGAGATCCCTAATACT 1207
Db 1153 ATTCCTTTAAATACAAAGTCATAGTTAAUUTTGGGCTATTGGAGAGATCCAAATAT 1212
QY 1208 TGGTCGGAACCTCGAAAGTTTAAACCAGAAAGGTTTAAAGATTGTCGAATTAATATAA 1267
Db 1213 TGGGATGATGAGAAAGCTTTTAAGCCTGAGAGATTTGAACATTAACCTCTTTGAATTTTGTCT 1272
QY 1268 GGGAGCAGATTTGAACTGGTACCATTGGTGCAGGAAAAAGAAATATGCTCGGCAATACT 1327
Db 1273 GGCATTAATTTGAATATCTTCTTTGGTAGTGAAGAGGATTTGCCCGGAAATATCA 1332
QY 1328 TCAGCTATTACCAATTTTGGAGTATGTCATTAATAATCTATTATATCATTTTAAATGGGAA 1387
Db 1333 TTTGTTTAGCTAATGTGTATCATCCATGGCTCAATTTGTTGTCATCTTCGATTTGGAGA 1392
QY 1388 CTGCGCGATGGAATTTACACCTCAACACTTGATATGATGATGAGCTATTGGCGGTGCTCTC 1447
Db 1393 CTTCCTACTGGGTGCGACCCAAATGACCTTGAATTTGACT-AGTTAGCTGGAGTAACACT 1451
QY 1448 AGGAAAAAATAGATCTTAAGTTGATTCCTATTCCATAT 1486
Db 1452 GGTAGGAAAGAGACCTTTACTTGTATTTCTACTCTCTTAT 1490

RESULT 6
US-08-948-564-3
; Sequence 3, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminezky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1691 base pairs
```

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..1545
; US-08-948-564-3

Query Match 17.8%; Score 308; DB 3; Length 1691;
Best Local Similarity 53.9%; Pred. No. 4.6e-73;
Matches 769; Conservative 0; Mismatches 620; Indels 39; Gaps 5;

QY 113 CTCCAGGCGCCATGGAAGTTTCCTCATAGAGTAATCTTCCTCATTTTACTCACTTCT 172
DB 145 CCCCCAGGACCAAGGACACTACCTCTCATAGGAAACATACACCAAGATTGTGGC---TCA 201

QY 173 GATCTAGGCCCATGAACGTTTATAGAGCCTTGGCTCAAAATTTATGAGCCTGTATAGTCTT 232
DB 202 CTGCCGGTTTCATTAATAAATTTGGCAGATAAGTATGTTCCATTTAATGCATCTA 261

QY 233 CAAATTTGGCCAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAGCAGCCAAAGAGTTATG 292
DB 262 AAACATAGGAGAGTGTCCAAACATCATAGTCACTTCCCCAGAAATGGCCCAAGAGATTAG 321

QY 293 AAAAATCAGGCTGATGCCCTTTCGCCCAACGCCCTATCGTCTTGGACGCACAGATTGTGTTT 352
DB 322 AAGACACATGATCTCAACTTCTCTGATAGCCAGACTTTGTATTGTCTAGATAGTTTCT 381

QY 353 TATAATCGGAAGATGTCTTTGTTTGTCTCATATGGAGATCACTGGAGGCAGATGAAGAAA 412
DB 382 TACAACGGTTCTGGCAATTTCTTTCAGTCAACATGGAGACTATTGGAGGCAACTAAGAAAG 441

QY 413 ATTTGATATCTGAAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
DB 442 ATATGCACAGTAGAGTTACTAACAGCAAAAGCGGTGCGAGTCTTTTCGGTCCATAAGAGAA 501

QY 473 GAAGAAATGGAGGATGCCATCACATCTCTCCGTTTCA-----AAGCGGATCT 520
DB 502 GAGGAGTGGCAGACTAGTTTAAATAATAGTGCNACTGCAAGTAGAAGAGGGGGTCC 561

QY 521 CCGGTCAATATTACAAAGATCAATTTATGGCATATTAATTTGCATCATATGAAGACATCC 580
DB 562 ATTTTAATCTCACCAGAGCATTTACTCAATGACTTTTGGGATAGCGGCACAGCGGT 621

QY 581 GTTGTGAATTTGAACAAAGAGATTTGCTGAGTGTTCGATGTCGATGCAATGAGCA 640
DB 622 TTTGTAAAGAGCAGATACCAACAAGTGTTCATATCAACATGCAATAAACAATGTATG 681

QY 641 GCGACGAGTTTGGCACCGCAGACGCTTTTCGACGTGGAAATTTACTTCACTATATCAT 700
DB 682 CTCTGGGAGGTTTCTGTGCTGATCTCTATCTCTCTAGTAGAGTGTTCATATGATG 741

QY 701 GGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACATATATCTTGAAGAGATT 760
DB 742 GGGGCGACGGGAAAATTGAAAAGTGCATAGATGACAGATAGGTGTGCAAGACATC 801

QY 761 CTTAATGACACAAAG-----CCAAATAGCCTTTTGAAGCGGATCACTTA 805
DB 802 ATCGACGAGCAAAAATAAGAAACAGAGCAGCGAGGCGGTGAAGCAGTGAAGATCTA 861

QY 806 ATGGATGTTTCTAATGAATCTTCAAAAATAATGGAACGTTTCCAGTGCAGTGAACAAAGAA 865
DB 862 GTTGATGTTCTCTCAAGTTTCAAAAGGATCGGA-----ATTTCGCTTGAATGATGAC 915

QY 866 AGCATCAAGACATCCGTTTGGCAATGTTTATGCGCGGAGCGAAACAACTTCGAAAGCT 925
DB 916 AACATTAAGCGGTCAATCAGGACATATTTCAATTTGGTGGAGGCGAAACATCATCTTCTGT 975

QY 926 ACAGATGGGTAATGGCAGAGCTGATGAAATCCAACTGAATTAAGAAAGCACAAGAA 985
DB 976 GTGGAAATGGGGATGTGAGATTGATAGAAACCCGAGGGTGTGGAAGAAAGCAAGCA 1035

QY 986 GAAGTTAGACAAAGTATTTGGTGAATGGCAAGTTGATGAATCAAGATTTTCATGATTTG 1045
DB 1036 GAGGTGAGAGAGTGTATATAGCAGGATATGTGGATGAGACAAATTTGCACCAATTTG 1095

QY 1046 AAATTTCTCAAGTTAGTGTATAAGAAACTCTAAGATTATCATCTCTCCGGT---TGTCTTG 1102
DB 1096 ATATACTTTAAAGTCCATCATCAAGAAACCATGAGGTTACATCCACTCTGCGCATTTGTTA 1155

QY 1103 ATTCCGAGGAGTGTAGAGAAACACCAAAATTTGATGGATATGAATTTTCATCCGAACACT 1162
DB 1156 GTTCTCTAGAGTAGTAGAGAAAGGTGCCAAATCAATGGATATGAGATACCCCTCTAAGACT 1215

QY 1163 CGAATTTGTGTAATGCTTTGGCGGATAGGAAGAGATCTCTAAATACTTTGGTCGGAACCTGGA 1222
DB 1216 AGATCATATTATCAATGCTTTGGGCAATTTGGAAGGAATCTTAAGTATTTGGGTGAACTGAG 1275

QY 1223 AAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTAAGGGAGCAACATTTGAA 1282
DB 1276 AGTTTAAACCTGAGAGGTTTCTTAATAGTCTCCATTTGATTTTAGGGGCACAGACTTTGAA 1335

QY 1283 CTGTTACCAATTTGGTCAGGAAAGAAATATGTCCTGGGCATTTACTTCAGCTATATCAAT 1342
DB 1336 TTTATCCCCATTTGGTCTGGAAGGAGGATCTCCCCGSCATTTACATTTGCCATACCCAAC 1395

QY 1343 TTGGAGTATGTCATTATTAATCTATTATATCATTTTAAATTTGGAACTGGCCGATGGAATT 1402
DB 1396 ATTGAGTTGCCACTTGTCTCAGTTACTTTACCATTTTGAATGGAGCTTCCCAATAAATG 1455

QY 1403 ACACCTCAAAACATTTGATATGATGAAAGCTATTTGGCGGTCTCTCAGGAAAAAATAGAT 1462
DB 1456 AAGAATGAAGAACTTGACATGCGGAGTCAATGGAATTTACTTTACGAAGACAAATGAC 1515

QY 1463 CTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTCAATATT 1510
DB 1516 CTCTGCTTGATTTCCCATCTACTCGTCTACCTTAAATGTATGAACAAT 1563

RESULT 7
US-09-292-768-5
; Sequence 5, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wbur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1494)
US-09-292-768-5

Query Match 14.6%; Score 253.2; DB 3; Length 1515;
Best Local Similarity 50.5%; Pred. No. 2.9e-58;
Matches 671; Conservative 0; Mismatches 648; Indels 9; Gaps 2;

QY 199 CTGGCTCAAAATTTATGACCTGTTATGAGTCTTCAAAATTTGCCCAAGTTTTCAGCTGTTCT 258
DB 186 CGTGGCAAGGATGACGCCCCCGTGGCCCACTGACGCTGGGTGAGGTGTTCTCCGTCGT 245

QY 259 CATTTCTTTCAGCTGAAGCAGCAAGAGGTTTATGAAAACTCAGGCTGATGCTTCGCCCA 318

Db 246 CTTTCTGTCGGGAGGCGACGAAAGGAGCGATGAAGCTGTGTAGACCCGGCGTGCAGAA 305
Qy 319 AGCCCTATCGTCTTGGAGCGACAGATTGTGTTTTATTAATCGGAAGAGTGTCTTGTTCG 378
Db 306 CGGTTTCAGAGCATCGGACGAGATCATGTGTGACGACACGAGGACATCATCTTCAG 365
Qy 379 TTCTATCGAGATCACTCGAGCGAGATGAAGAAATTTGGATCTTGAATTTCTGAGTGC 438
Db 366 CCCCTACAGCGAGCACTCGGCGCAGATCGCAAGATCTCGCTCTCCGAGCTCTCTCCTC 425
Qy 439 CAAAAAGTTCAATCCTCCAGGTTAATCGAGAGGAAGAAATGGAGGATGCCATCACATT 498
Db 426 CGCAACGTCCGCTCTTCGGCTTCATCCGAGGAGCAGAGGTCTCGCGCTCTCTCCGCCA 485
Qy 499 CTTCCGTTTCAAGCGGATCTCCCGTCAATATTAACAAAGATCATTTATGCAATTATAAT 558
Db 486 CTTCCGCTCTTCGAGGCGGCGCGTGGACATGACGAGAGGATAGAGACGCTGACCTG 545
Qy 559 TTCGATCATGATAAGAAATCCGTTGGTAAATGTAA---GCAAAAGAAAGATTCCTGAG 615
Db 546 CTCCATCATCTCAGGGGGCGCTTCGGGAGCGTGTACAGGACAAACGCGGAGCTGTGGG 605
Qy 616 TGTTCGATGAGTCAATGAGCGAGCGAGGTTTTCGACCGCAGACGCTTTTCCGAC 675
Db 606 GCTGTCAAGGACGCGCTCAGCATGGCTCGGGGTTCGAGCTCGCGGACATGTTCCCTC 665
Qy 676 GTGAAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAGCGCTTTCATCAGGA 735
Db 666 CTCGAGCTCTCAACCTCTCTGCGNAACAGAGCAAGCTCTGGAGGATGCGCGCG 725
Qy 736 GATTCAGCATATCTTGAAGAGATCTTAAAGAAACAAAGCAATAAG-----CCTTT 789
Db 726 CGTGACACCATCTCGAGGCGCATCGTCAAGCAGCAAGTTTCAAGAAAGCGCGGAGTT 785
Qy 790 TGAAGCGATACTTAATGAGTGTCTTATGATCTTCAAAAAATGNAAGTTCAGT 849
Db 786 CGGCGGAGGAGCATCATCGAGTCTCTTCAAGATGAGAGGCGCCAGCATCAAGT 845
Qy 850 GCCAGTGACAAAGCAAGCATCAAAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGGA 909
Db 846 CCCCATCAACCACTCCATCAAGCCTTCATCTCGATAGTTCTCAGCAGGAGCTGA 905
Qy 910 AACATCTCGAAGCTACAGATGGGTAAATGCGAGAGCTGATGAAGAAATCCAACTGAAT 969
Db 906 GACATCTCAACACCACTTATGAGTGTGCGGAGCTGATGAGGAAACCCGCGAGTGA 965
Qy 970 AAGAAAGCACAAAGAGTTAGACAAAGTATTTGTTGTAATGGGAAAGTTGATGAATC 1029
Db 966 GCGAAAGCGGAGGAGTGAAGCGGCTGAGGAGAGACGAACTGGGAGCTGA 1025
Qy 1030 AAGATTTTCATGATTTGAAATCTTCAAGTTAGTGTAAAGAACTCTTAAGATTAATCC 1089
Db 1026 TGATGTGCAAGAGCTTAAGTACATGAATCGGTGTGAAGGAGACGATGAGGATGACCC 1085
Qy 1090 TCCGTTGTCTGATTCGAGGAGTGTAGAGAAACACACAAATGATGATGATGAAT 1149
Db 1086 TCCGATCCGTTGATCCCGAGATCATGACAGAAAGATTCGTTGTTAAGCGGTATACGAT 1145
Qy 1150 TCATCCGACACTCGAATTTGTTGTAATGCTTTGGCGATGAGGAGATCCTTAATCTTG 1209
Db 1146 TCGAACAGGCGAGATCATGATCAACGCTGTGTCATGGGAGGAACTCTCTCTACTG 1205
Qy 1210 GTCCGAACCTGGAAAGTTTAAACCGAGAAAGTTTAAAGATTTGCAATTTGATTAAGG 1269
Db 1206 GGAAAAACCGATACCTTTTGGCCCGAAAGTTTGAACCAAGTTTCAAAGGATTTTCATGG 1265
Qy 1270 GAGGACATTTGAGTGGTACCATTTGGTCAGAGAAAGAAATATGCTCGGCAATTAATTC 1329
Db 1266 AAATGATTTTCGATGTTGCTCCGTTTCGAGCGGAGAAAGATCTGCCCGGCTTGAATTT 1325
Qy 1330 AGCTATTACCAATTTGGAGTGTGATTAATTAATCTATTATATCATTTTAAATTTGGAATC 1389
Db 1326 CGGTCGCGAAAGTTGAGGTTCCATTTGCGGAGCTTCTTTACCACTTCGATCGAAGTT 1385

Qy 1390 GGCGATGGAATTACACCTCAACACTGATATGACTGAAGCTATTGGCGGTGCTCTCAG 1449
Db 1386 GGCGAAGGAATGAACCTTCTGATATGGACATGTCTGAGCGGAAGCCTTACCGGAAT 1445
Qy 1450 GAAAAAATAGATCTTAAGTTGATTCCTTATTTCCATATCAAGTTAGCTTAGGCTCAATAT 1509
Db 1446 ACTAAGAACAAATCTTCTTCTTGTTCACACCCCTACGATCCTCTCATGATCAATCAT 1505
Qy 1510 TTCTTGAT 1517
Db 1506 CTTTGTCT 1513

RESULT 8

US-08-881-784-5
; Sequence 5, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESSEE: PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881.784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1762 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mentha spicata
; INDIVIDUAL ISOLATE: cDNA encoding
; INDIVIDUAL ISOLATE:
; IMMEDIATE SOURCE:
; CLONE: pSM12.2
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 558..1212
; OTHER INFORMATION: /product= "Probe LH-1 (Figure 4A)"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 39..538
; OTHER INFORMATION: /product= "Probe LH-2 (Figure 4A)"
US-08-881-784-5
Query Match 14.2%; Score 245.4; DB 3; Length 1762;

Best Local Similarity 50.5%; Pred. No. 4e-56;		Matches 706; Conservative 0; Mismatches 676; Indels 15; Gaps 4;	
Qy	102	AACAGATCCACCTCCAGGCGCATGGAAGTTCCTATCATAGTAACTCTTCCTCATTTAT	161
Db	111	AACAAAACCTACTCTCGAGCCCTCGAAGCTCGCGGTGATCGGCACCTCC-----ACT	164
Qy	162	TACTCATTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTG	221
Db	165	TCCTGTGGGAGGCTTCCCGAGCAGCTGTTTAGGAGCATGCCAGAGTACGGCCGG	224
Qy	222	TTATGAGTCTCAAAATGGCCAAAGTTTCAGCTGTTGTCTATTTCTCAGCTGAAGACCA	281
Db	225	TGGCCACATGTCAGCTGGGAGAGTGTACTCGGTGGTGTCTCGCGGAGGAGCGA	284
Qy	282	AAGAGTTATGAAAACCTCAGGCTGATCGCTTCGCCCAACGCCCTATCGTCTTGGACGAC	341
Db	285	ACGAGCGATGAAGTGTGGACCCGAATTCGCCGCCGTTTCAGCGCATCGGGTCCA	344
Qy	342	AGATTGTGTTTTAATCGGAAGATGTCTTGTGCTTCATATGGAGATCACTGGAGGC	401
Db	345	GGACCATGTGTACGACAAAGATGACATCATCTTCAGCCCTTACACGATCACTGGGCC	404
Qy	402	AGATGAAGAAAATTTGGATACATTGAATTTCTGAGTGCCAAAAGTTTCAATCCTCAGGT	461
Db	405	AGATCGGAGGATCTCGGTGACAGAGCTGCTGAGCCCGAAGAACGTCAGGTCTTCCGGT	464
Qy	462	TAATCCGAGAGGAAGATGGAGGATGCCATCACATTCCTCGTTCGAAAGCCGATCTC	521
Db	465	ACATAAGGAGGAGATCGAGCGCCTCATCCGGCTGCTCGGGTCTCGGGGGAGCGC	524
Qy	522	CGGTCAATATTACAAAGATCAATTTATGGCATTATAATTTTCGATCATGATAAGAACATCG	581
Db	525	CGGTGACGTGACGGAGGAGTGTGCAAGATGCTGTGTCTGTCGTGTCAGGGCGGCT	584
Qy	582	TTGTGTAATGTAAGCAAAAGAAAGATGTCG- -AGTGTGCGGATGAGTCAAT- -GAGG	638
Db	585	TGGGAGTGTGTCGAAGGACCAAGGTTGCTTGGCGAGTTGGTGAAGAGTCGCTGGCAT	644
Qy	639	CAGGACAGATTTTGGCCCGCAGACGCTTTTCGACGCGGAATTACTTCACTATATCA	698
Db	645	TGGCTTCGGGTTTGAGCTGGCGGATCTCTACCTTCTCTCATGGCTCTCAACCTGCTTA	704
Qy	699	TTGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATCTTTGAAGAGA	758
Db	705	GCTTGAACAGTACAGTTGACAGAGATGGCGCGCGCTCGATCACATCTTGTATGGT	764
Qy	759	TTCTTAATGAACAAAGCAATAAG- ------CTTTTGAAGCGGATAACTTTAATGGATG	812
Db	765	TCCTGGAGGATAGGGAGAAAGAGAGCGCGGAGTTTGGAGCGGAGACATCGTCGACG	824
Qy	813	TTCTATGATCTTCAAAAATGNAAGGTTCCAGTGCCAGTGCAGAACGAAACATCA	872
Db	825	TTCTTTTCAGGATGACAGAGGCGCAGACATCAAAATTTCCCATTTACTTCCAAATTCATCA	884
Qy	873	AAGCATCCGCTTTTGAATGTTTTACTGCGGAGCGAAACAACTTCGAAAGCTACAGAT	932
Db	885	AGGGTTTCATTTTCGACACCTTCTCGCGGAGAGTGAAACGTTCTTGACGACCATCTCAT	944
Qy	933	GGGTAAATGGCAGCTGATGAAAAATTCAACTGAACTAAGAAAAACAAGAAAGATTA	992
Db	945	GGGCGTTGTGCGAACTGATGAGGAATCCGGCGAAGATGGCCAAAGTGGCGGCGAGTAA	1004
Qy	993	GACAGTATTTGGTGAATGGGAAAAGTTGATGAATCAAGATTTTATGATTTGAATTTCT	1052
Db	1005	GAGAGCGCTCAAGGGAAGACAGCTCGTGGATTTGAGCGAGGTGCAAGAGCTTAAATATACC	1064
Qy	1053	TCAAGTGTAGTGTAAAGAAACTCTAAGATTACATCTCCGCTGCTGTGTTGATTCCGAGGG	1112
Db	1065	TCAGATCGGTGTAAAGGAGACTCTGAGGCTGCACCTTCCCTTTTCCATTAATCCCAAGAC	1124
Qy	1113	AGTGTAGAGAAACAAACGAATTTGATGATATGAAATTCATCCGAACACTCGAATTTGTG	1172

Db	1125	AATCAGGGAAGAATTCGAGGTTAACGGGTACACGATTCGGGCCAAAACCTAGAACTTCTCA	1184
Qy	1173	TGAATGCTTGGCGGATAGGAAGATCCTAACTACTTGGTCGGAACCTGGAAGTTTAAACC	1232
Db	1185	TCAGCTCTGGGCTATCGGAAGGATCCCAATACTTGGGAAGATCCCGACACCTTCGGCC	1244
Qy	1233	CAGAAAGGTTTAAAGATTGTGCAATTGTGCAATTGATTATAAAGGGACGACATTTGAACTGGTACCAT	1292
Db	1245	CTGAGAGATTTCGATGAGTTTCCAGGGAATTTTCATGGGAACGATTTTCGAGTTTCATCCCAT	1304
Qy	1293	TTGTGTCAGGAAAAAAGATATGTCCTCGCATTACTTTCAGCTATTACCAATTTGGAGATG	1352
Db	1305	TGGGGCGGGTGAAGAATCTGCCCGGTTTACATTTCCGGCTGGCAAAATGTTGAGATCC	1364
Qy	1353	TCATTATAAATCTATTATATATATTAAATTTGGAAGTGGCCGATGGAATTTACACCTCATA	1412
Db	1365	CATTGGCGCACTGCTTACCACCTTCGACTGGAATTTGCCAAGGAATGACTGATGCGG	1424
Qy	1413	CATTGTATATGACTGAAGCTATTGGCGGTCTCTCAGGAAAAAAATAGATCTTAAAGTTGA	1472
Db	1425	ACTTGGACATGACGGAGACCCCGAGTCTTCTGGGCCAAAAGAAATGTTTGCTTGG	1484
Qy	1473	TTCTATTTCATATCAAA 1489	
Db	1485	TTCCCACTCTATAAAA 1501	

RESULT 9
US-09-292-768-1
; Sequence 1, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wsl13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1762
; TYPE: DNA
; ORGANISM: Mentha spicata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20) .. (1507)
US-09-292-768-1

Query Match 14.2%; Score 245.4; DB 3; Length 1762;
Best Local Similarity 50.5%; Pred. No. 4e-56;
Matches 706; Conservative 0; Mismatches 676; Indels 15; Gaps 4;

Qy	102	AACAGATCCACCTCCAGGCGCATGGAAGTTCCTATCATAGTAACTCTTCCTCATTTAT	161
Db	111	AACAAAACCTACTCTCGAGCCCTCGAAGCTCGCGGTGATCGGCACCTCC-----ACT	164
Qy	162	TACTCATTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTG	221
Db	165	TCCTGTGGGAGGCTTCCCGAGCAGCTGTTTAGGAGCATGCCAGAGTACGGCCGG	224
Qy	222	TTATGAGTCTTCAAAATGGCCAAAGTTTCAGCTGTTGTCTTTCAGCTGAAGACCA	281
Db	225	TGGCGACGTGTCAGCTGGGAGAGTGTACTCGGTGGTGTCTGTCGTCGGCGGAGGACGA	284
Qy	282	AAGAGTTATGAAAACCTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGAC	341
Db	285	ACGAGCGATGAAGTGTGTCGACCCGAATTTCCGCCGACCGGTTTCGACCGCATCGGGTCCA	344


```
Db 465 ACATAAGCAGGAGAGATCGAGCGCTCATCGGCTGCTCGGTGCTCGGGGGAGCGC 524
QY 522 CGGTCAATATTACAAAGCATATTATGGCATTATAATTCGATCATGAAGAATCCG 581
Db 525 CGGTGACGTGACGAGAGAGGTGCGAGATGTCGTGTGTCGTGTCAGGGCGGCGT 584
QY 582 TTGGTAATTGTAAGCAAAAGAAAGATTGCTG--AGTGTTCGCCGATGCACTCAAT--GAGG 638
Db 585 TCGGGAGTGTGTCGAAGACACAGGTTTCGTTGGCGAGTTGGTGAAGGAGTCTCGTGCAT 644
QY 639 CAGCGACGAGTTTGGCACCGCAGACGCTTTTCGAGCTGGGAAATTAATTCATATATCA 698
Db 645 TGGCGTCCGGTTTGAAGTGGCGGATCTCTACCTTCTCATGGCTCTCAACCTGCTTA 704
QY 699 TTGGAGCTGAGTCAAAACCCAGCGTTTGCATCAGGAGATTGACGATATATCTTTGAAGAGA 758
Db 705 GCTTGAACAGTACAGGTTTCAGAGGATGCGCGCTCGATCACATCTTTGATGGGT 764
QY 759 TTCTTAATGAACACAAAGCCAAATAAG-----CCTTTGAAGCGGATAACTTAATGGATG 812
Db 765 TCCTGGAGGACATAGGAGAGAGAGCGCGAGTTTGGAGCGGAGCATCGTCGACG 824
QY 813 TTCTATTGAATCTTCAAAAATAATGGAACGTTCCAGTGCAGTGAACAAAGCAAGCATCA 872
Db 825 TTCTTTTCAGGATGCAAGAGGCGGACATCAAAATTTCCATTACTTCCAAATTGCATCA 884
QY 873 AAGCATCCGTTTTCGAAATGTTTACTCGCGGAGCGAACAACCTTCGAAAGCTACAGAAT 932
Db 885 AGGTTTCAATTTTCGACACCTTCTCCGCGGAGCTGAAAGCTTTCGAGACCATCTCAT 944
QY 933 GGTAATGCGAGAGCTGATGAAATAATCCAACCTAAGAAAGCAGCAAGAAAGTTA 992
Db 945 GGGCGTTGTCGAACATGATGAGGAATCCGCGAAGATGGCAAGTTCGAGCGGAGTAA 1004
QY 993 GACAGTATTGTTGTAATGGGAAAGATTGATGAATCAAGATTTTCATGATTTGAAATCT 1052
Db 1005 GAGAGCGCTCAAGGGAAGACAGTCTGTGATTTGAGCGAGTGCAGAGCTAAATAAC 1064
QY 1053 TCAAGTTAGTGTAAAGAACTCTAAGATTACATCTCGGTGCTTTCGATTCGAGG 1112
Db 1065 TGAGATCGGTGTTAAAGGAGACTCTGAGGCTGCACCTTCCCTTTCATTAATCCCAAGAC 1124
QY 1113 AGTGTAGAGAAACACACGAAATTTGATGGATATGAAATTCATCCGAACACTCGAATTTGTG 1172
Db 1125 AATCCAGGAGAGATCCGAGGTTAACGGGTACACGATTTCGGCCAAATCTAGATCTTCA 1184
QY 1173 TGAATCCTTGGCGATAGGAAGATCTTAATCTTTGGTGGAACTCGGAAAGTTTAAACC 1232
Db 1185 TCAACGCTCTGGCTATCGGAAGGATCCCAATATCTGGGAAGATCCCGACACCTTCCGCC 1244
QY 1233 CAGAAAGTTTAAAGATTGCAATTTGATATATAAGGAGACATTTGAACTGGTACCAT 1292
Db 1245 CTGAGAGATTCGATGAGGTTTCAGGGAATTCATGGGAACGATTTTCGAGTTTCATCCCAT 1304
QY 1293 TTGGTGCAGGAAAGAAATATGCTTCCTGGCATTACTTCAGCTATTACCAATTTGGAGTATG 1352
Db 1305 TCGGGCGGTCGAGAACTCTGCCCGGTTTACATTTTCGGGTGGCAATTTGATATCC 1364
QY 1353 TCATTATAATCTATTATATCAATTTTAAATGGGAAGTGGCCGATGGAATTAACCTCAAA 1412
Db 1365 CATTTGGCGCAACTGCTCTACCACTTCGATGGAATTTGCACAAAGGAATGACTGATGCCG 1424
QY 1413 CACTTGATATGACTGAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTTAAGTTGA 1472
Db 1425 ACTTGACATGACGGAGACCCAGAGTCTTTCTGGGCCAAAAAAGAAAAATGTTGCTGG 1484
QY 1473 TTCTTATTCATATCAAA 1489
Db 1485 TTCCACACATCTATAAA 1501
```

RESULT 11

US-09-292-768-67

```
; Sequence 67, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wbur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; NAME/KEY: misc feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
; US-09-292-768-67
```

Query Match 14.1%; Score 243.8; DB 3; Length 1665;

Best Local Similarity 50.2%; Pred. No. 1.1e-55;

Matches 692; Conservative 0; Mismatches 672; Indels 15; Gaps 3;

```
QY 199 CTTGCTCAAAATTTATGGACCTGTTATGAGTCTTCAATTTGCCAAGTTTCAGCTTTGT 258
Db 204 CGTGCGAAGCAGTACGGCCCACTGGCGCACGTCGAGCTCGCGAGGTGTTCTCCGTCGT 263
QY 259 CATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCCATCACATT 318
Db 264 GCTCTGTCCTCCGCGAGGCCACGAAAGGCGCATGAGCTGGAGCCCGCTTCGCGGA 323
QY 319 AGCCCTATCTGCTTTGGACGACAGATTGTGTTTTATATCGGAAAGATGTCTTGTTCG 378
Db 324 CCGGTTTCGAGAGCATCGGGACGAAGATCATGTGTTACGACACGACATCATCTTCAG 383
QY 379 TTCATATGGAGATCACTGGAGCAGATGAGAAATTTGGATACATTGAATTTCTGAGTGC 438
Db 384 CCCCTACAGCGTGCATCGGCCAGATCGGGAAGATCTGCGTCTCCGAGCTCCTCAGCGC 443
QY 439 CAAAAAAGTTCAATCTCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATT 498
Db 444 CGCGAAGCTCGCTCTCTTCGGCTTCATCAGCGACGACGAGGTGTCGCGCTCTCGGCCA 503
QY 499 CTTCCGTTTCGAAAGCC-----GGATCTCCGCTCAATATTACAAAGATCAATTTATGGCAT 552
Db 504 CTTCCGCTCTCTCGCGCGCGGGGGGCGCGTGGACCTCAGGAGCGGATAGCGACGCT 563
QY 553 TATATTTTCGATCATGATAAGAACATCCGTTGGTAAATTGTAA--GCAAAAAGAAAGATT 609
Db 564 GACGTGCTCCATCATCTCGAGGGCGCGTTCGGGAGCGGTGATCAGGGACCAAGAGAGCT 623
QY 610 GCTGAGTGTTCGGATGCAATGAGGACGACGAGTTTGTGGCACGCGAGAGCTTT 669
Db 624 GGTGAGCTGGTGAAGGACGCCCTCAGCATGCGCTCCGGTTCGAGCTCCGCGCATGTT 683
QY 670 TCCGACGTGGAAATTTACTTACTATATCATTTGAGAGCTGAGTCAAAACCCAGCGTTTGA 729
Db 684 CCCCTCTCCAGCTCTCTCACTTCTGCTGGAACAAGCAAGCTGTGGAGGATGCG 743
```


Db 885 AGGTTTCATTTTCGACACCTTTCGCGGGAGCTGAAACGCTTCGACGACCATCTCAT 944
QY 933 GGTATATGCGAGCTGATGAAAAATCCAACTAAGAAAGCACAAGAGAGTTA 992
Db 945 GGGCTTGTGCGAATGATGAGGATCCGCGGAGATGCGGAGGAGGTAA 1004
QY 993 GACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTTGAATTTCT 1052
Db 1005 GAGAGCGCTCAAGGGAAGACAGCTGCTGATTTGAGCGAGGTGCAAGAGCTAAAATACC 1064
QY 1053 TCAAGTTAGTGTAAAGAACTTAAGATTACATCTCGGTGTGCTTGTGATTCGAGGG 1112
Db 1065 TGAGATCGGTGTAAAGGAGACTGAGGCTGACCCCTCCCTTCCATTAATCCCAAGAC 1124
QY 1113 AGTGTAGAGAAACACACGAAATTTGATGGATATGAAATTCATCCGAACACTCGAAATTTGTG 1172
Db 1125 AATCCAGGAGAAATCGGAGTTTACGGGTACAGATTTCGGGCAAACTAGAAATCTTCA 1184
QY 1173 TGAATGCTTTGGCGATPAGAAAGATCTTAATATCTTTGGTTCGGAACCTGGAAGTTTAAAC 1232
Db 1185 TCAAGCTCTGGCTATCGGAAGGATCCCAATACTTGGGAAGATCCCGACACCTTCGCGC 1244
QY 1233 CAGAAGGTTTAAAGATTGTCGAATTTGATTAATAAGGAGACATTTGAACCTGGTACCAT 1292
Db 1245 CTGAGAGATTCGATGAGGTTTCCAGGGATTTTCATGGGAAACGATTTTCGAGTTTCATCCCAT 1304
QY 1293 TTGTGTCAGGAAAGAAATATGTCCTGCGATTACTTTCAGCTATTACCAATTTGGAGTAG 1352
Db 1305 TCGGGCGGTGCAAGAAATCTGCCCGGTTTACATTTTCGGGTGCGAAATGTTGAGATCC 1364
QY 1353 TCATTATATAATCTATTATATCAATTTTAAATGGGAACCTGCGCGATGGAATTTACACCTCAA 1412
Db 1365 CATTGGCGCACTGTCTTACCACTTCGACTGGAATTTGCCAAGGAATGACTGATGCGC 1424
QY 1413 CACTTGATATGACTGAAGCTATTGGGGTGCTCTCAGGAAAAAATAGATCTTAAGTTGA 1472
Db 1425 ACTTGACATGACGAGACCCAGGTCTTTCTGGGCCAAAAAAGAAAAATGTTTGTGG 1484
QY 1473 TTCTATTCCATATCAA 1489
Db 1485 TTCCACACTCTATAAA 1501

RESULT 13

US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESSES: PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.

; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mentha x piperita
; IMMEDIATE SOURCE:
; CLONE: pPM17
; US-08-881-784-8

Query Match 14.0%; Score 242.2; DB 3; Length 1665;
Best Local Similarity 50.1%; Pred. No. 2,9e-55;
Matches 691; Conservative 0; Mismatches 673; Indels 15; Gaps 3;
QY 199 CTTGGCTCAAAATTTATGACCTGTTATGAGTCTTCAAATTTGCGCAAGTTTTCAGCTGTTGT 258
Db 204 CGTGGCAAGCAGTACGGCCAGTCGCGCACGTCGAGCTCGGCGAGGTGTTCTCCGCTG 263
QY 259 CATTTCCTTCAGCTGAAGCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCCCTTCGCCCA 318
Db 264 GCTCTGTCGTCGCGAGGCGCACGAAGAGGCGATGAAGCTGGTGGAGCCCGCTTCGCGGA 323
QY 319 AGCGCTTATCGTCTTGGAGCGCACAGATTGTTGTTTTATATCGGAAAGATGTCCTGTTTGC 378
Db 324 CGGTTTCGAGGACATCGGAGCAGATCATGTGTTACGACACGACGACATCATCTTCAG 383
QY 379 TTCATATGGAGATCATCTGAGGACAGATGAAGAAAAATTTGGATCTTTGAATTTCTGAGTGC 438
Db 384 CCCCTACAGCGTGCACTGGCGCCAGATCGGAAGATCTCGCTCTCGAGCTCCTCAGCGC 443
QY 439 CAAAAAAGTTCATCTCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATT 498
Db 444 CGCAACGTCGCTCTTCGGCTTCATCAGGACGAGCAGGTTGCCGCTCCTCGGCCA 503
QY 499 CCTCGTTTCGAAAGCC-----GGATCTCCGGTCAATATTACAAAGATCATTTATGGCAT 552
Db 504 CCTCGCTCTCCGCGCGCGGGGAGGCGCTGACCTCAGGAGCGGATAGCGACGCT 563
QY 553 TATAATTTTCGATCATGATAAGAACATCCGTTGGTGAATTTGTAA---GCAAAAAGAAAGATT 609
Db 564 GACGTGCTCCATCATCTGACGGCGCGGTTTCGGGAGCGTGATCAGGGACACGAGGAGCT 623
QY 610 GCTGAGTGTTCGGATGCAATGAGGCGACGAGTTTGGCACCAGAGCGCTTT 669
Db 624 GGTGAGCTGTGAAGGACGCCCTCAGCATGGCGTCCGGTTTCGAGCTCGCCGACATGTT 683
QY 670 TCCGAGTGGAAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAGCGTTTGA 729
Db 684 CCCCTCTCCAAAGCTCCTCACTTCTCTGCTGGAAACAAGAGCAGCTGTGGAGGATGCG 743
QY 730 TCAGGAGATTGACGATATCTTTGAAGAGATTTCTTAATGAACACAAAGCAATAAG----- 784
Db 744 CGCGCGCTCGACGCCATCTCGAGGCCATCGTGGAGGAGCACAAGCTCAAGAAGAGCGG 803
QY 785 -CCTTTTGAAGCGGATACTTAATGATGTTCTTATTGATCTTCAAAAAATGGAAACGT 843
Db 804 CGAGTTTGGCGCGAGGACATTTATGACGTACTCTTTAGGATGCAAGAGGATAGCCAGAT 863
QY 844 TCCAGTGGCAGTGAACAAAGCAAGCATCAAAAGCATCCGTTTTCGCAAAATGTTTACTGCCGG 903
Db 864 CAAAGTCCCATCACCACCAACGCCATCAAGCCCTTCATCTTCGACACGTTCTCAGCGGG 923
QY 904 GAGCGAAACAACTTCGAAAGCTACAGATGGGTAAATGGCAGAGCTGATGAAAAATCCAAAC 963
Db 924 GACCGAGACATCATCAACCAACCCCTGTGGGTGATGGCGGAGCTGATGAGGAATCCAGA 983

[illegible]

RESULT 14

```

US-09-292-768-3
; Sequence 3, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wbur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19) .. (1518)
US-09-292-768-3

```

Query Match
14.0%; Score 242.2; DB 3; Length 1665;

Query Match	100%	Score 214.2; E-35	Length 2003
Best Local Similarity	50.1%	Pred. No. 2.9e-55;	
Matches 691; Conservative	0; Mismatches 673;	Indels 15; Gaps	38

QY 199 CTTGGCTCAAATTTATGGACCTGTTATGAGTCTTCAAATTGGCCAAGTTTCAGCTGTTGT 258

Db 1284 CATGGGAACGATTCGAGTTCATCCATTTGGAGCTGGAAGAAGAACTCTGCCCGGTTT 1343
QY 1324 TACTTCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATATATTTAATG 1383
Db 1344 GAATTTCCGGTTGGCAATTTGAGTCCATTTGGCAGACGCTTCTTTACCACTTCAGCTG 1403
QY 1384 GGAACCTGGCGATGGAATACACCTCAACACCTTGATATGACTGAAGCTATTGGCGGTGC 1443
Db 1404 GAAGTTGGCGGAAGGAATGAACCTTCCGATATGGACATGTCTGAGCAGAAAGCCCTTAC 1463
QY 1444 TCTCAGGAAGAAATAGATCTTAAGTTGATTCCTATTCCATATCAATGATGAGCTC 1503
Db 1464 CGGAATAAGAAAGAAACATCTTACTCGTTCCACACCTAGATCCTTCTCTCATGATC 1523
QY 1504 AATATTTCTTGATTAACATAGGAGGTGGAATATATATAATAAATCTTTAATTTAAGAT 1562
Db 1524 AATATATCTTTAATTTGCTCTTTGAATAAAGAGTGCATATATATATATATATATAT 1582

RESULT 15
US-09-292-768-69
; Sequence 69, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase protein variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-69

Query Match 14.0%; Score 242.2; DB 3; Length 1665;
Best Local Similarity 50.1%; Pred. No. 2,9e-55;
Matches 691; Conservative 0; Mismatches 673; Indels 15; Gaps 3;

QY 199 CTTGGCTCAATTTATGACCTGTTATGAGTCTTCAATTTGGCAGATTTTCAAGTGTGTG 258
Db 204 CGTGGCAAGCAGTATCGGCCCAAGTGGCGCACGTCGAGCTCGGCGAGGTGTTCTCGGTCT 263
QY 259 CATTTCTTCAGCTGAAGCAGCAAGAGGTATGAAAATCTCAGGCTGATGCCCTTCGCCCA 318
Db 264 GCTCTGTCCTCCCGAGGCCAGAGAGCGCATGAGCTGGTGGACCCGCGCTTCGCGGA 323
QY 319 AGCCCTCATGCTTTGGAGCAGATTTGTGTTTATATATCGGAAGATGTTCTGTTTGC 378
Db 324 CCGGTTCCGAGAGCATCGGACGCAAGATCATGTGGTACGACACGACGACATCATCTTCAG 383
QY 379 TTCATATGAGATCACTGAGGAGCAGATGAAGAAATTTGGATACATTTGATTTCTGAGTGC 438
Db 384 CCCCTACAGCGTGCACCTGGCGCCAGATGCGGAAGATCTGCGTCTCCGAGCTCCTCAGCGC 443

QY 439 CAABAAGTTCAATCCTCCAGGTTAATCCGAGAGGAGAAATGGAGGATGCCATCACATT 498
Db 444 CCGCAACGTCGCTCTTTTCGGCTTTCATCAGGAGGAGGAGGTGTCCTCCGCTCTCTCGGCCA 503
QY 499 CTTCCGTTTGGAAAGCC-----GGATCTCCGGTCAATATTACAAAGATCATTTTATGGCAT 552
Db 504 CTTCCGCTCTCTCGGCGCGCGGGGAGGCGCTGGACCTCACGGAGCGGATACGGACGCT 563
QY 553 TATTAATTTTCGATCATGATTAAGAACATCCGTTGGTAAATTTGTAA---GCAAAAAAGAAAGATT 609
Db 564 GAGTGTCTCATCATCTGAGGCGCGGCTTCGGAGCGTGCATCAGGAGACCAACGAGAGCT 623
QY 610 GCTGAGTGTTCGCGATGACATGAGGACGACGAGCTTTTGGCACCGCACACGCTTT 669
Db 624 GGTGAGCTGTTGAGGAGCGCTCAGCATGGCGTCCGGGTTTCGAGCTCGCCACATGTT 683
QY 670 TCCGAGCTGGAAATTAATTTCACTATATATATGAGCTGAGTCAAAACCCAGGCGTTTGCA 729
Db 684 CCCCTCTCCAAAGCTCTCAACTTGTCTGTGCTGGAACAGCAAGCTGTGGAGGATGCG 743
QY 730 TCAGGAGATTGACGATATATCTTTGAAGAGATTCTTAAATGAACACAAAGCCAAATAG----- 784
Db 744 CCGCGCGCTCGACGCGCATCTCTGAGGCCATCTCTGAGGCCATCTCTGAGGAGCAAGCTCAAGAAGAGCGG 803
QY 785 -CCTTTTGAAGCGGATACTTAATGAGTCTTCTTATGATCTTCAAAAAATGGAACCT 843
Db 804 CAGATTTGGCGGAGGACATTAATGACGTACTCTTTAGGATGCAAGAGGATAGCAGAT 863
QY 844 TCCAGTGGCAGTGACAAACGAAAGCATCAAAAGCATCCGTTTTCGAAATTTTACTCCCGG 903
Db 864 CAAAGTCCCATCACCAACGCGCATCAAAAGCTTCACTCTCGACAGCTTCTCAGCGG 923
QY 904 GAGCGAAACAACTTCGAAAGCTACAGATGGGTATGGCAGAGCTGATGAAAAATCCAAC 963
Db 924 GACCGAGACATCATCAACCCACCCCTGTGGGTGATGGCGGAGCTGATGAGGAATCCAGA 983
QY 964 TGAACATAAGAAAACACAAAGAAAGTTAGACAGTATTGTTGTAATCGGAAAGCTTGA 1023
Db 984 GGTGATGGGAAAGCGCAGCGAGCTGAGCGCGCTGAAGGGGAAACGACATGGGA 1043
QY 1024 TGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATT 1083
Db 1044 CGTGACGACGTCGAGGAGCTTAAGTACATGAAATCGGTGGTGAAGGAGACCATGAGGAT 1103
QY 1084 ACATCTCTCGGTTGCTTGTATTCGAGGAGGTGTAGAGAAACACACAGATTTGATGGATA 1143
Db 1104 GCACCTCCGATCCCGTTGATCCCGAGATCATGSCAGAGAAGAAATCGAGGTCAACCGGTA 1163
QY 1144 TGAATTCATCCGAACACTCGAATTTGTTGTAATCTTGGCGGATAGGAAGAGATCCTAA 1203
Db 1164 CACGATTCGAAATAAGCCAGATCATGATCACTGTTGTTCCATGTTGGTGAAGATCTCT 1223
QY 1204 TACTTGGTCCGAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTTGTGCAATTTGATTA 1263
Db 1224 CTACTGGGAAAAACCCGAGACCTTTTGGCCCGAAAGTTTGAACCAAGTCTCGAGGATTT 1283
QY 1264 TAAAGGAGACATTTTGAATCTGATACCATTTGGTTCAGGAAAGAAATATGTCCTGGCAT 1323
Db 1284 CATGGAAACGATTTTCGAGTTTCATCCCATTTGGAGCTGGAAGAAATCTGCCCGGTTT 1343
QY 1324 TACTTCAGCTATTACCAATTTGGAGTATGTCATTAATAATCTATTATATCATTTTAAATG 1383
Db 1344 GAATTTCCGGTTGGCAATTTGAGTCCCATTTGGCAGCTCTTTTACCACCTTCGACTG 1403
QY 1384 GGAATCGCGATGGAATTTACACCTCAAAACACTTGTATGATGACTGAAGCTATTGGCGGTGC 1443
Db 1404 GAAAGTTGCGGAAGGAAATGAACCTTTCCGATATGACATGTCTGAGGAGAGAGGCTTAC 1463
QY 1444 TCTCAGGAAAGAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTC 1503
Db 1464 CGGAATAAGAAAGAAACAAATCTTACTCTGTTCCCAACCCCTACGATCTCTCTCATGATC 1523

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:03:08 ; Search time 8489 Seconds
(without alignments)
5649.731 Million cell updates/sec

Title: US-10-759-813-2

Perfect score: 2574

Sequence: 1 MEQKLSFPISILISFLVLI.....KIDKLIPYQVSLGNSIS 500

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/abs/ABSWEB_spool/US10759813/runat_31052006_060304_2592/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05h
-USER=US10759813 @CN_1_1_5548@runat_31052006_060304_2592 -NCPU=6 -ICPU=3
-NO_MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: gb_vi.*
11: gb_ov.*
12: gb_hg.*
13: gb_in.*
14: gb_om.*
15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Match Length	Description
1	2574	100.0	AX366075 Sequence
2	2574	100.0	AF406732 Euphorbia
3	1164	45.2	CS223906 Sequence

4	1149	44.6	1673	2	CS223964	Sequence
5	1149	44.6	1673	2	CS223984	Sequence
6	1148	44.6	1673	2	CS223962	Sequence
7	1144	44.4	1857	2	CS223948	Sequence
8	1142	44.4	1673	2	CS223966	Sequence
9	1142	44.4	1701	4	STU296346	Solanum t
10	1131	43.9	1610	2	CS223950	Sequence
11	1131	43.9	1610	2	CS224014	Sequence
12	1129.5	43.9	1576	2	CS223902	Sequence
13	1128.5	43.8	1566	2	CS223944	Sequence
14	1128.5	43.8	1581	2	CS223904	Sequence
15	1128.5	43.8	1664	2	CS223952	Sequence
16	1127.5	43.8	1781	2	E63735	Cytochrome
17	1127.5	43.8	1781	2	AR205821	Sequence
18	1127.5	43.8	1781	2	AF122821	Capsicum
19	1122.5	43.6	1784	2	CS249182	Sequence
20	1119	43.5	1665	2	A29004	pCGP147 DNA
21	1119	43.5	1665	2	I47318	Sequence 29
22	1106.5	43.0	1509	4	AB028462	Petunia x
23	1102	42.8	99089	4	AP007265	Lotus cor
24	1100	42.7	1515	4	AF368376	Nicotiana
25	1097.5	42.6	1572	2	CS223930	Sequence
26	1088	42.3	1666	2	I28313	Sequence 24
27	1088	42.3	1666	2	CS224040	Sequence
28	1082.5	42.1	1559	2	Y10493	G. max mRNA
29	1071.5	41.6	1800	4	GM450CP7	CR954196
30	1068	41.5	132419	4	CR954196	Medicago
31	1059.5	41.2	1754	4	GM450CP3	Y10490 G. max mRNA
32	1056.5	41.0	1789	4	DQ335785	Medicago
33	1054.5	41.0	116998	4	AC140774	Medicago
34	1052.5	40.9	1641	4	SCU48434	U48434 Solanum cha
35	1045.5	40.6	1641	4	AF000403	Lotus jap
36	1042.5	40.5	1882	4	CR0238612	Catharant
37	1039.5	40.4	1506	4	AY596975	Centauriu
38	1036	40.2	3107	4	SCU48435	U48435 Solanum cha
39	1026.5	39.9	122392	4	AC137079	Medicago
40	1026	39.9	110957	12	CT009652	Medicago
41	1026	39.9	122965	4	CR932040	Medicago
42	1023	39.7	139522	4	AC145061	Medicago
43	1023	39.7	259855	12	AC145024	Medicago
44	1022.5	39.7	1564	2	CS223914	Sequence
45	1019	39.6	122965	4	CR932040	Medicago

ALIGNMENTS

RESULT 1	AX366075	Sequence 1	1733 bp	DNA	linear	PAT 15-FEB-2002
LOCUS	AX366075	Sequence 1 from Patent WO208269.				
DEFINITION	AX366075					
ACCESSION	AX366075					
VERSION	AX366075.1	GI:18697502				
KEYWORDS						
SOURCE	Euphorbia lagascae					
ORGANISM	Euphorbia lagascae					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;					
	rosids; eurosids I; Malpighiales; Euphorbiales; Euphorbiaceae; Euphorbioideae;					
	Euphorbiae; Euphorbia.					
REFERENCE	1					
AUTHORS	Cahoon, E.B.					
TITLE	A cytochrome p450 enzyme associated with the synthesis of _g(d)?12					
JOURNAL	-epoxy groups in fatty acids of plants					
	Patent: WO 0208269-A 1 31-JAN-2002;					
	E. I. du Pont de Nemours and Company (US)					
FEATURES	Location/Qualifiers					
source	1..1733					
	/organism="Euphorbia lagascae"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:54672"					
ORIGIN						
Alignment Scores:						

Pred. No.:	1.11e-286	Length:	1733
Score:	2574.00	Matches:	500
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-10-759-813-2 (1-500) x AX366075 (1-1733)			
Qy	1	MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle	20
Db	14	ATGGACGAGAAAATCTCTTTTCGGAGCATTTAATAAGTTTCTGCTTGTGTTTAATC	73
Qy	21	LeuValValValMetArgLeuTrpLysLysGlnAsnProProGlyProTrpLysPhe	40
Db	74	TTAGTAGTAGTCATGAGTGTGTGAAGAGAAACAGAAATCCACCCTCCAGGCGCATGGAAGTTT	133
Qy	41	ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe	60
Db	134	CCATCATGAGTAATCTTCCCTCAATTAATTACTACTTCTGATCTAGGCCAAGAAGCTTTT	193
Qy	61	ArgAlaLeuAlaGlnIleTyrglyProValMetSerLeuGlnIleGlyGlnValSerAla	80
Db	194	AGAGCCTTGGCICAAATTTATGACCTGTATTAGTCTTCAAAATGGCCAAAGTTTCAGCT	253
Qy	81	ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe	100
Db	254	GTTGTCAATTTCTTTCAGCTGAAGCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCCTTC	313
Qy	101	AlaGluArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu	120
Db	314	GCCCAAGCCCTATCTCTTGGACGCACAGATGTGTTTTATAATCGAAAGATGCTTG	373
Qy	121	PheAlaSerTyrglyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu	140
Db	374	TTTGCTTCATATGGAGTCACTGGAGGCAGATGAAGAAATTTGGATACTTGAATTTCTG	433
Qy	141	SerAlaLysLysValGlnSerSerArgIleuIleArgGluGluGluMetGluAspAlaIle	160
Db	434	AGTGCCAAAAAAGTTCAATCTCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATC	493
Qy	161	ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrglyIle	180
Db	494	ACATTTCTCGTTCGAAAGCCGGATCTCCGGTCAATATTACAAGATCATTTATGGCAAT	553
Qy	181	IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu	200
Db	554	ATAATTTTCGATCATGATAAGAACATCCGTTGGTAAATCTGAAGCAAAAAGAAAGATTGCTG	613
Qy	201	SerValAlaAspAlaValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPhePro	220
Db	614	AGTGTTGCCGATCGACTCAATGAGCGCGCAGGTTTTTGGCACCGCAGACGCTTTTCOG	673
Qy	221	ThrTrpLysLeuLeuHisIleTyrlleIleGlyAlaGluSerLysProArgArgLeuHisGln	240
Db	674	ACGTGGAATTAATCTTCACTATATCATTTGGAGCTGAGTCAAAACCAGCGGTTTGATCAG	733
Qy	241	GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLysProPheGlu	260
Db	734	GAGATTGACGATATCTTCAAGAGATTCTTAATGAACACAAAGCCAATAAGCCTTTTGAA	793
Qy	261	AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro	280
Db	794	GCGGATAACTTAATGATGTGTTCTATTGAATCTTCAAAAATAATGAAAACGTTTCCAGTGCCA	853
Qy	281	ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr	300
Db	854	GTGACAAACGAAAGCATCAAGCATCCGTTTGCAATGTTTACTGCGCGGAGCGGAACA	913
Qy	301	ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg	320
Db	914	ACTTCGAAGCTACAAATGGGTAAATGGCAGAGCTGATGAAATATCAACTGAACATGAAGA	973

/db_xref="GI:18157659"
 /translation="MEQKNLSPPSILISFLVLIVVMRLWKQMPPPGPKFPFIIG
 NPLHLTSLDLGHERFRAAOIYGPVMSIQIGVSAVVISAEAAKEVWKMTQADAPAO
 RPIVLDAQIYFVNRKDQVLFASGDHWRQMKKIWLEFLSAKKVQSSRLURESEMEDAI
 TFLRSKASPVNITKIIYGIISIMRTSVGNCKQKERLUSVADAVNEAATSGTADA
 FPTWKLHYSIIIGAESKPRRLHQEIDIDILEILNEHKANKPFADNLMDVLLNLRKNGM
 VFPVTNESIKASVQMFTAGSETTSKATEWMAELMKNPTELKRAQEVRQVFGEMG
 KYDESFRDLKFTFLVVKETLRHLPPVLIIPRECRETLTIDGVEIHPNTRIIVNNAWAI
 GRDPNTWSPPGKFNPERFKCAIDYKTTFFELVPFGAGKRIICPGITSALTINLEYVIN
 LLYHFNWELADGITPTQDWTBTAIGGALKKIDKLIPIPYQVSLGNSIS"

ORIGIN

Alignment Scores:

Pred. No.: 1,13e-286 Length: 1758
 Score: 2574.00 Matches: 500
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-759-813-2 (1-500) x AF406732 (1-1758)

Qy	1	MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle	20
Db	21	ATGGAGCAGAAAATCTCTTTCCGAGCATTTTAAATAGTTTTCGCTGTGTTTAAATC	80
Qy	21	LeuValValValMetArgLeuTrpLysLysGlnAsnProProGlyProTrpLysPhe	40
Db	81	TTAGTAGTAGTCATGAGGTGTGGAAAGAACAGATCCACCTCCAGGGCCATGGAAAGTTT	140
Qy	41	ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe	60
Db	141	CCTATCATAGGTAATCTCTCATTTTACTTACTCACCTTCGTATGATAGGCCATGAACGTTT	200
Qy	61	ArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAla	80
Db	201	AGAGCCTTCGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGGCCAAAGTTTCAGCT	260
Qy	81	ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe	100
Db	261	GTGTGATTTCTTCAGCTGAAGCAGCCAAAGAGTTTATGAAACTCAGGCTGATGCCCTTC	320
Qy	101	AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu	120
Db	321	GCCCAAGGCCCTATCGCTTTCAGCGCACAGATTTGTGTTTATAATCGGAAGATGCTCTG	380
Qy	121	PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu	140
Db	381	TTTGCTTCATATGGAGATCACCTGGAGGCAGATGAAAGAAATTTGGATCTTGAATTTCTG	440
Qy	141	SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIle	160
Db	441	AGTGCCAAAAGATTCATCTCCAGTTTAAATCCGAGAGAGAAATGGAGAGTCCATC	500
Qy	161	ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIle	180
Db	501	ACATTCTCCGTTTGGAAAGCCGATCTCCGGTCAATATTACAAAGATCATTTATGCAATT	560
Qy	181	IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu	200
Db	561	ATAATTTTCGATCATGATAGAACATCCGTTGGTAATTTGTAAGCAAAAGAAAGATTGCTG	620
Qy	201	SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro	220
Db	621	AGTGTGCCGATGCAGTCAATGAGGCGCGACAGTTTTTGGCACCCGACAGCCTTTCCG	680
Qy	221	ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysPheProArgArgLeuHisGln	240
Db	681	ACGTGGAAATTACTTCACTATATCATTTGAGCTGAGTCAAAACCCAGCGGCTTTGCATCAG	740
Qy	241	GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLysProPheGlu	260
Db	741	GAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCCAATAAGCCTTTGAA	800

Qy	261	AlaAspAsnLeuMetAspValLeuLeuAenLeuGlnLysAsnGlyAsnValProValPro	280
Db	801	GGGATAACTTAATGATGTTCTTATTGAATCTTTCAAAAATATGGAACGTTCCAGGTGCCA	860
Qy	281	ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr	300
Db	861	GTGCAAAACGAAAGCATCAAGCATCCGTTTTTGCAAAATGTTTACTCCCGGAGCGAAACA	920
Qy	301	ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg	320
Db	921	ACTTCGAAAGCTACAGAATGGGTAATGGCAGAGCTGATGAAAATCAACTGAACTAAGA	980
Qy	321	LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg	340
Db	981	AAAGCACAGAGAGTAGATGATTTTGGTGAATATGGAAAAGTTGATGATGATCAAGA	1040
Qy	341	PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro	360
Db	1041	TTTTCATGATTGAAATTTCTCAAGTTAGTGGTTAAAGAAATCTTAAGATTTACATCTCCG	1100
Qy	361	ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis	380
Db	1101	GTGTGCTTGATTCGAGGGAGTGTAGAGAAACAACAAGATTCGATGATGATGAAATTCAT	1160
Qy	381	ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer	400
Db	1161	CCGACACCTCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCTTAATCTTGGTGC	1220
Qy	401	GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr	420
Db	1221	GAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTTGTCAATTCATTATAAGGGACG	1280
Qy	421	ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla	440
Db	1281	ACATTGAACCTGATACCACTTTCGTCGAGGAAAGAAATATGCTGCGCATTTACTTCAGCT	1340
Qy	441	IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla	460
Db	1341	ATTACCAATTTGGAGTATGTCATTATAATCTATTATATCATNTTTAATTTGGGNACTGGCC	1400
Qy	461	AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys	480
Db	1401	GATGAAATTTACACCTCAACACTTCGATATGACTGAAGCTATTGGCGGTCTCTCAGGAAA	1460
Qy	481	LysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeuGlySerAsnIleSer	500
Db	1461	AAAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTCAAAATATTTCT	1520
RESULT 3			
LOCUS	CS223906	1538 bp	DNA
DEFINITION	Sequence 162 from Patent WO2005111217.		
ACCESSION	CS223906		
VERSION	CS223906.1		
KEYWORDS	GI:83685434		
SOURCE	Nicotiana tabacum (common tobacco)		
ORGANISM	Nicotiana tabacum		
REFERENCE	1 Xu D.		
AUTHORS	Nicotiana nucleic acid molecules and uses thereof		
TITLE	Patent: WO 2005111217-A 162 24-NOV-2005;		
JOURNAL	U.S. Smokeless Tobacco Company (US)		
FEATURES	Location/Qualifiers		
source	1. 1538		
	/organism="Nicotiana tabacum"		
	/db_type="unassigned DNA"		
	/db_xref="taxon:4097"		
ORIGIN			

Qy 12 LeuIleSarPheLeuValLeuValLeuValMetArgLeuTrpLysLysGln 31
 |||:::|| ||| :: :|||::|||
Qy 12 LeuIleSarPheLeuValLeuValLeuValMetArgLeuTrpLysLysGln 31
 |||:::|| ||| :: :|||::|||

Qy	378	GlulleHisProAsnThrArgAlleValValAsnAlaTrpAlaIleGlyArgAspProAsn	397
Db	1188	ACTATTCTGTGAAGACCAAAAGTCATGGTTAAATGTTTGGGCTTTGGGAAGAGATCCAAAA	1247
Qy	398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
Db	1248	TATTGGAAATGACGCGAGAAACTTTTATGCCAGAGAGATTTCGACGAGTCTCTAAGGATTTT	1307
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Db	1308	GTTTGGTAATAATTTTGAATATCTCCATTTGGTGGCGAAGGAGGATTTCCTGGGATT	1367
Qy	438	ThrSerAlaIleThrAsnLeuGluTyrValIleAsnLeuLeuTyrHisPheAsnTyr	457
Db	1368	TCGTTTGGCTTAGCTAATAGCTTATTTGGCAATGGCTCAATTACTATATCATCTTCGATTGG	1427
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
Db	1428	AAATCCTCTGCTGGAATCGAACCAAGCGACTTGACTGGTTGAGTTGGTGGAGTAACT	1487
Qy	478	LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln	492
Db	1488	GCCGCTAGAAAAAGTGACCTTTACTTGGTTCGCACTCCTTATCAA	1532
RESULT 5			
LOCUS	CS223984	1673 bp	DNA
DEFINITION	Sequence 240 from Patent WO2005111217.		
ACCESSION	CS223984		
VERSION	CS223984.1	GI:83695871	
KEYWORDS			
SOURCE	Nicotiana tabacum (common tobacco)		
ORGANISM	Nicotiana tabacum		
REFERENCE 1			
AUTHORS	Ku,D.		
TITLE	Nicotiana nucleic acid molecules and uses thereof		
JOURNAL	Patent: WO 200511217-A 240 24-NOV-2005;		
FEATURES	U.S. Smokeless Tobacco Company (US)		
source	Location/Qualifiers		
ORIGIN	1..1673		
	/organism="Nicotiana tabacum"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4097"		
Alignment Scores:			
Pred. No.:	5,7e-122	Length:	1673
Score:	1149.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	44.6%	Indels:	16
DB:	2	Gaps:	5
US-10-759-813-2 (1-500) x CS223984 (1-1673)			
Qy	12	LeuIleSerPheLeuLeuValLeuLeuLeuValValMetArgLeuTrpLysGln	31
Db	54	TTGGTTTCCATTTTCCTATTCTATCTTCCCTCTTTTGTAAAGATATGGAAGACTCC	113
Qy	32	AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu	46
Db	114	AATAGCCAAAGCAAAAAGTTGCCACCAAGTCCATGGAAACTACCAATACTAGGAAGTATG	173
Qy	47	ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	174	CTTCATATGGTT-----GGTGGACTACCAACCATGCTCTTAGAGATTTCAGCCAAAAA	227
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerAla	86

```
Db 228 TATGACCACCTTATGACCTTCAATTAGGTGAAGTTTCTGCGGTGTGGTACTTCTCTCT 287
Qy 87 GluAlaLaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 288 GATACGGCAAGAAGATATAAATACTCATGACATCGCTTTTGGCTAGCGCTAGCCTT 347
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 348 TTGGCCCGGAGATTGTCTGTACAAATAGGTCTGATCTAGCCTTTTGGCCCTATGGCGAC 407
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 408 TATTGGAGACAAATCGTAAATAATGTGTCTTGGAAAGTCTCAGTGCACCAAGAATGTCGG 467
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 468 ACATATTAGCTCTATTAGCGGAATGAAGTTCTTCCTCTCATTAATTTATCGGTCATCT 527
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 528 TCTGTTGAACCTATTAAATGTTACGGAAGGATCTTTTGTTCACAAAGCTCCATGACATGT 587
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 588 AGATCAGCGTTTGGCAAGTGTTCAAAGACGAAGCAAAATTATACAACTAATTAAAGAA 647
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 648 GTGACTACTTTAGCAGGAGGGTTTCATGTGGCTGACATATTCCTTCACTGAAGTTTCTT 707
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspIle 245
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 708 CATGTGCTCAGTGGGAATGAAGGTAAAGATTATGAATGCACACCAATAAGATAGATCCCAT 767
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaasnLysPro 258
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 768 GTTGAGAATGTCATCAATGAGCACAAGAAAATCTTGCATTTGGGAAAACCTAATGGACG 827
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 259 PheGluAlaAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 828 TTAGGAGGTGAAGATTAAATGATGTCTTCTTAAGACTTATGAATGATGGAGCGCTTCAA 887
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 888 TTTCTATATCACCACGACCAACATCAAGCCATAAATTTTGTGACATGTTTGTGCTGCCGGGACA 947
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 948 GAGACTTTCATCTCAACAAATTTGTGGCTATGGTAGAAATGGTGAAATAATCCAGCCGTA 1007
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1008 TTCGCGAAGCTCAAGCAGAAGTAAGAGAAGCATTTAGAGGAAAAGAAACTTTTCGATGAA 1067
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 339 SerArgPheHisAspLeuLysPheLysLeuValValLysGluThrLeuArgLeuHis 358
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1068 AATGATGTGGAGGAGCTAAACTACCTAAAGTTAGTAATAAAGAAAGAACTCTAAGACTTCAT 1127
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 359 ProProVal---ValIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1128 CCACCGGTTCCACTTTTCTCCCAAGAGAAATGTAGGAGAGACAAATATAAAGCGCTAC 1187
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1188 ACTATTCTCTGTAAGACCAAGCTCATGTTAATGTTTGGGCTTTTGGGAAGAGATCCAAAA 1247
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1248 TATTGGAATGACCGAGAACTTTTATGCCAGAGAGATTTGACGAGTGTCTCTAAGGATTTT 1307
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 418 LysGlyThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1308 GTTGGTAATATTTTGAATATCTTCCATTTTGTGGCGGAGAGGATTTGTCTGGGATT 1367
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1368 TCGTTTGGCTTAGCTAATGCTTATTGGCTCAATTACTATATCACTTCACTTCGATTGG 1427
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1428 AAACCTCCCTGCTGGAATCGAAACCAAGCAGCTTGGACTTGAAGTGTGGGTAAGTACT 1487
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1488 GCCCTAGAAAAAGTACCTTTACTTGTTCGACTCCCTTATCAA 1532
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
RESULT 6
CS223962 1673 bp DNA linear PAT 15-DEC-2005
LOCUS Sequence 218 from Patent WO2005111217.
DEFINITION CS223962
ACCESSION CS223962
VERSION CS223962.1 GI:83695860
KEYWORDS
SOURCE
ORGANISM Nicotiana tabacum (common tobacco)
            Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE
1 Xu,D.
  AUTHORS Nicotiana nucleic acid molecules and uses thereof
  TITLE Patent: WO 200511217-A 218 24-NOV-2005;
  JOURNAL U.S. Smokeless Tobacco Company (US)
FEATURES
            source
            1..1673
            /organism="Nicotiana tabacum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4097"
ORIGIN
Alignment Scores:
Pred. No.: 7,44e-122 Length: 1673
Score: 1148.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.8% Indels: 16
DB: 2 Gaps: 5
US-10-759-813-2 (1-500) x CS223962 (1-1673)
Qy 12 LeuIleSerPheLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 54 CTGGTTTCCATTTTCTTATTTCTTCTCTCTTTTGTAGGGTATGGAGAACTCC 113
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 32 AsnPro-----ProGlyProTrpLysPheProIleIleGlyAsnLeu 46
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 114 AATAGCCAAACAAAAAGTTGCCACCAAGTCCATGTAACCTACCAATACTAGGAAGTATG 173
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 47 ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 174 CTTCATATGGTT-----GGTGGACTACCAACCATGTCCTTAGAGATTTAGCCAAAAA 227
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 228 TATGACCACTTATGACCTTCAATTAGGTGAAGTTTCTGCGGTGTGGTACTTCTCT 287
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 87 GluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 288 GATACGCAAAAGAAAGTATTAAAAAATCATGACATCGCTTTTGGCTCTAGGCGTAGCCTT 347
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 348 TTGGCCCGGAGATTGTCTGTACAAATAGGTCTGATCTAGCCTTTTGGCCCTTATGGCGAC 407
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Db 408 TATTGGAGACAAATCGCTAAATATATGCTTTGGAAAGTGTCTCAGTGCCCAAGAAATGTTGCG 467
 QY 147 SerSerArgLeuLeuArgGluGluMetGluAaspAlaIleThrPheLeuArgSerLys 166
 Db 468 ACATTAGCTCTATTAGCGGAATGAAGTCTTCGTCTCATTAATTTATCCGGTCTCT 527
 QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrglyIleIleSerIleMetIle 186
 Db 528 TCTGTGTGAACCTATTATATGTCAGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
 QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAaspAla 205
 Db 588 AGATACGCTTTGGGCAAGTGTCAAAGACGAACAAATTTATCAACATTAATAAGAA 647
 QY 206 ValAsnGluAlaIleThrSerPheGlyThrAlaAaspAlaPheProThrTrpLysLeuLeu 225
 Db 648 GTGATACTCTTAGCAGGAGGTTTGATGTGGCTGACATATTCCTTCATCGAAGTTCTT 707
 QY 226 HisTyrlleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAaspIle 245
 Db 708 CATGTGCTCAGTGAATGAAGGTAAAGATTATGAATGCACACCATAAGGTAGTGCATT 767
 QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
 Db 768 GTTGAGAATGTATCAATCAGCACAGAAATAATCTTTGCAATTTGGGAAATACTAATGGAGCG 827
 QY 259 PheGluAlaAaspAsnLeuMetAaspValLeuLeuAsnGlnLysAsnGlyAsnValPro 278
 Db 828 TTAGGAGGTGAAGATTATTAATGATGTTCTTAAGACTTATGAATGATGGAGCCTTCAA 887
 QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
 Db 888 TTTTCTTATCACCACCAACATCAAGCTATAATTTTGGACATGTTTCTGCTCGCGGACA 947
 QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
 Db 948 GAGACTTCATCTGCTCAACAAATTTGTGGGCTATGGTAGAATGGTGAATAATCCAGCGTA 1007
 QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAaspGlu 338
 Db 1008 TTCGGGAAGCTTCAGCAGAAGTAGAGAAAGCATTTAGAGGAAGAAACTTTTCGATGAA 1067
 QY 339 SerArgPheHisAaspLeuLysPheLysLeuValValLysGluThrLeuArgLeuHis 358
 Db 1068 AATGATGTGAGGAGCTAACTACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTCAT 1127
 QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAaspGlyTyr 377
 Db 1128 CCACCGGTTCACCTTTGCTCCCAAGAGAAATGTAGGGAAGAGACAAATATNAACGCTAC 1187
 QY 378 GluIleHisProAsnThrArgIleValValAsnAlaIleAlaIleGlyArgAaspProAsn 397
 Db 1188 ACTATTCCTGTAAAGACCAAGCTCATGTTATATGTTGGGCTTTGGGAAGAGATCCAAA 1247
 QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAaspCysAlaIleAaspTyr 417
 Db 1248 TATTGGAATGACGACAGAACTTTTATGCCAGAGAGATTGAGCAGTGCTCTAAGGATTTT 1307
 QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
 Db 1308 GTTGGTAATAATTTTGAATATCTCCATTTGTTGGCGGAAGGAGGATTTGCTCTGGGAT 1367
 QY 438 ThrSerAlaIleThrAsnLeuGluTyrrValIleIleAsnLeuLeuTyrrHisPheAsnTrp 457
 Db 1368 TGGTTTGGCTTAGCTTAATGCTTATTGGCATTTGGCTCAATTAATATATCATCTTCGATGG 1427
 QY 458 GluLeuAlaAaspGlyIleThrProGlnThrLeuAaspMetThrGluAlaIleGlyAla 477
 Db 1428 AAATCCCTGCTGGAATCAACCAAGCGACTTGGACTTGCATGAGTTGGTTGGAGTAAC 1487
 QY 478 LeuArgLysLysAlaAspLeuLysLeuIleProIleProTyrrGln 492
 Db 1488 GCCGCTAGAAAAAGTGACCTTTTACTTGGTTGGGACTCTCTTATCAA 1532

RESULT 7

CS223948
 LOCUS 1657 bp DNA linear PAT 15-DEC-2005
 DEFINITION Sequence 204 from Patent WO2005111217.
 ACCESSION CS223948
 VERSION 1 GI:83685853
 KEYWORDS
 SOURCE Nicotiana tabacum (common tobacco)
 ORGANISM Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiids; Solanales; Solanaceae; Nicotiana.

REFERENCE

1 Xu,D.
 Nicotiana nucleic acid molecules and uses thereof
 Patent: WO 2005111217-A 204 24-NOV-2005;
 U.S. Smokeless Tobacco Company (US)
 Location/Qualifiers
 1. 1657
 /organism="Nicotiana tabacum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4097"

ORIGIN

Alignment Scores:

Pred. No.: 2,13e-121 Length: 1657
 Score: 1144.00 Matches: 223
 Percent Similarity: 66.5% Conservative: 107
 Best Local Similarity: 45.0% Mismatches: 152
 Query Match: 44.4% Indels: 14
 DB: 2 Gaps: 6

US-10-759-813-2 (1-500) x CS223948 (1-1657)

QY 10 SerIleLeuIleSerPheLeuLeuValLeu---IleLeuValValMetArgLeuTrp 28
 Db 37 TCTAACTTAGTTGCTATCTTCTCTCTCTCAGCATCTTCTTCTTATTTCAAAAAATGG 96
 QY 29 Lys-----LysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
 Db 97 AAAACACAGAAAATAAATTTGCTCTCTGCTCCATGAAATTTACCTTTTATGGAAATTTA 156
 QY 47 ProHisLeuLeuLeuThrSerAaspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
 Db 157 CACCATTTGGCTGTGGCAGGTCCACTTCTCTACCATGGCTTAAAAAATTTAGCCAAACGC 216
 QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
 Db 217 TATGTCTCTCTTATGCAATTTACAACTTGGACAAATTTCTTACACTCGTCATCATCACCT 276
 QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAaspAlaPheAlaGlnArgProIleVal 106
 Db 277 CAAATGGCAAAAGAAAGTACTAAAAAATCTCAGACCTCGCTTTTGGCCACTAGACCAAGCTT 336
 QY 107 LeuAaspAlaGlnIleValPheTyrrAsnArgLysAaspValLeuPheAlaSerTyrrGlyAasp 126
 Db 337 GTCTGGCGGACATCATTCACCTACGACAGCAGCAGACATAGCATTTTCGCCATACGGTGA 396
 QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
 Db 397 TACTGGAGACAAATTCGTAATAATTTGCATATTTGCAACTCTTGAGTGCCCAAGATGGTCAAG 456
 QY 147 SerSerArgLeuIleArgGluGluMetGluAaspAlaIleThrPheLeuArgSerLys 166
 Db 457 TTTTGTAGCTCGATTCGCCAAGATGAGTCTCGAAGATGGTTTCATCTATATGACGACGAGC 516
 QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrglyIleIleSerIleMetIle 186
 Db 517 CCCAATCTTCCAGTCAATCTTACCACCAAGATTTTGTGGTTACGAGTTTCGTAATTTCT 576
 QY 187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAaspAla 205
 Db 187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAaspAla 205


```

Db      768 GTTGAGAAATGTCATCAATGAGCACAAGAAAAATCTTGCAATGGGAAAACTAATGAGCG 827
Qy      259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db      828 TTAGGAGGTGAAGATTAAATGATGTTCTCTAAGACTTATGAATGATGAGCGCTTCAA 887
Qy      279 ValProValThrAsnGluSerLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db      888 TTTCTCTATCACCACGACACATCAAGCTATAATATTTTGGACATGTTTCTGCGCGGAGC 947
Qy      299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db      948 GAGACTTCACTGCAACAATTTGTGGCTATGTTAGTAATGGTGAATAATCCAGCCGTA 1007
Qy      319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db      1008 TTCGCGAAGCTCAAGCAGAGTAGAAGAGCATTTAGAGGAAAAGAACTTTCGATGAA 1067
Qy      339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db      1068 AATGATGTGGAGGAGCTAAACTACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCAT 1127
Qy      359 ProProVal--ValleulleProArgGluCysArgGluThrThrArgileaspGlyTyr 377
Db      1128 CCACCGGTTCACCTTTTGTCTCCCAAGAGAAATGTAGGGAAGACAAATATAAACGGCTAC 1187
Qy      378 GluIleHisProAsnThrArgileValValAsnAlaTipAlaIleGlyArgAspProAsn 397
Db      1188 ACTATTCCTGTAAAGCCAAAGTCATGTTAATGTTGGCTTTGGGAAGAGATCCAAAA 1247
Qy      398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db      1248 TATTGGAATGACGCAGAAACTTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTT 1307
Qy      418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgileCysProGlyIle 437
Db      1308 GTTGGTAATAATTTTGAATATCTTCATTTGTGGCGGAAGAGGATTTGCTCGGATT 1367
Qy      438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db      1368 TCGTTGGCTTAGCTAATGCTTATTGTCATGGCTCAATTAATCACTATCACTTCGATTGG 1427
Qy      458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db      1428 AAACCTCCCTGCTGGAATCAACCAAGCGACTTGGACTTCACTGAGTTGGTTGGAAGTAACT 1487
Qy      478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db      1488 GCCGCTAGAAAAAGTGACCTTTTACTTGGTTGCGACTCCTTATCAA 1532

RESULT 9
STU296346
LOCUS      Solanum tuberosum mRNA for cytochrome P450 (CYP71D4 gene).
DEFINITION
ACCESSION  AJ296346
VERSION     AJ296346.1 GI:12331297
KEYWORDS   CYP71D4 gene; cytochrome P450.
SOURCE      Solanum tuberosum
ORGANISM    Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE
1
AUTHORS    Schneider M.
TITLE      Cloning of a potato gene (CYP71D4) induced after infection with
JOURNAL    Phytophthora infestans
REFERENCE  2 (bases 1 to 1701)
AUTHORS    Schneider M.
TITLE      Direct Submission
JOURNAL    Submitted (18-JAN-2001) Schneider M., Plant Biology, University of
Fribourg, 3, rue Albert Gockel, CH-1700 Fribourg, SWITZERLAND
FEATURES

```

```

1. .1701
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Datura"
/db_xref="taxon:4113"
1. .1701
/gene="CYP71D4"
23. .1531
/gene="CYP71D4"
/function="unknown"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAC24711.1"
/db_xref="GI:12331298"
/db_xref="GOA:O9AVQ2"
/db_xref="InterPro:IPR001128"
/db_xref="InterPro:IPR002401"
/db_xref="UniProtKB/TrEMBL:O9AVQ2"
/translation="NQFUSLASIFLFLSLFLLRKWNKNSQS KKLPPGPKWKLPLGSG
MLHMAGGLPHVLRDLAKKYGPLMHQLQGEVAVVVTSPDMAKEVLKTHDIAPSRPK
LLAPEICVNRSDIAPCPYGDYRWQRKICVLELLSAKVRSGIRRDGSDVRLVNFPI
RSSGEPVNFTEFLFTSSMTCSRAGKVPKEODKFIOLIKEVIGLGGDFVDLFFP
SLKFHLVSGMKGKIMNAHKHVDVAIVEDVINEHKKKPAIGKTNGALGGDLIDLVLRL
MNDGGLQPTITNDNIKAIIFDMAAGTETSSSTLVAMVOMKNSPSVIAKQAQAEVREA
FKDKETFDENDVEELKYLKVIKETLRHPVPLVPRECEETDINGTIPVTKVM
VNWALGRDPKYWDAAESFKPERFQRSVDFVGNPFYLPFGGRRICPGISPLGANV
YLPLAHLLYHFDWKLPIGMEPKDLNLTELGVTVAAKDDLLLVATPYEPFRQ"

ORIGIN
Alignment Scores:
Pred. No.:      3,76e-121      Length:      1701
Score:          1142.00        Matches:    217
Percent Similarity: 64.2%      Conservative: 101
Best Local Similarity: 43.8%    Mismatches: 161
Query Match:     44.4%        Indels:     16
DB:              4            Gaps:       5

US-10-759-813-2 (1-500) x STU296346 (1-1701)
Qy      12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysGln 31
Db      38 TTGGCTTCCATCTTCTTTCTATCTTTTCTTTTAAAGGAAATGGAAAACTCG 97
Qy      32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db      98 AATGCCAATCGAAAAAATGGCTCCAGTCCATGGAACTTCCTTACTAGGAAGTATG 157
Qy      47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db      158 CTTTCAT-----ATGGCTGGTGGACTTCCACACCATGTCTTAGAGATTAGCCAAAAA 211
Qy      67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db      212 TATGACCACCTATGATCATCTTCAACTGGTGAAGTCTCTGCAAGTTGTAGTAACCTTCCT 271
Qy      87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db      272 GATATGGCGAAGAAGTACTAAAACTCATGACATCGCTTCGCTCTAGGCTTAACTT 331
Qy      107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db      332 TTGGCCCGGAAATTTGTTTGTACCAACAGGTCTGACATTTGCCCTTGTGCCCTACGAGAT 391
Qy      127 HisTrpArgGlnMetLysIleIleTrpIleLeuGluPheLeuSerAlaLysValGln 146
Db      392 TACTGGACACAAATTCGTAATAATTTGTGTCTTGGAAATTTGTTGAGTCCCAAGATGTCGG 451
Qy      147 SerSerArgLeuIleArgGluGluMetGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db      452 TCATATGCTCGATTAGCGCGCATGAGTTGATCGCTTGTAAATTTTATCCGGTCATCT 511
Qy      167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186

```

```
Db 512 TCGGGTGAGCCGGTTAAATTTTACTGAAAGGTTGTTTTTGTTCACAAAGTTCAATGACATGT 571
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 572 AGATCAGCGTTCGGGAAAGTGTTCAAGACAGGACAAATTTATACAACTAATCAAGAA 631
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 632 GTGATTGGGTAGCAGGAGGATTTGATGGGTGATATCTTTCCCATCATCTGAAGTTTCTC 691
Qy 226 HisTyrIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle 245
Db 692 CATGTGCTTAGTGAATGAAGTAAATTTATGAACGCTCATCATAGGTAGATGCAATTT 751
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 752 GTTGAAGATGTCATCAATGAGCACAAAGAAATTTGCAATTTGGGAAACAAATGATGTCGA 811
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 812 TTAGGTGGTGAAGATCTAATGATGCTCTTAAAGACTTTATGAATGATGGAGGCCCTCAA 871
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 872 TTTCCGATCACCAACGACAACTAAGCTATTATTTTCGACATGTTTGTGTCAGGAACA 931
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 932 GAGACTTCATCGTCAACACTTGTCTGGCAATGGTGCAAAATGATGAAACCCCAAGTGTA 991
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 992 ATCGCAAGCTCAAGCAGCAAGTCGGAGAGCCTTTAAAGACCAAGAACCGTTCGATGAA 1051
Qy 339 SerArgPheHisAspLeuLysPheLysValLeuValLysGluThrLeuArgLeuHis 358
Db 1052 AATGATGTAGAGAGCTGAATACTTAAAGTTAGTCATTAAAGAACTCTAAGACTCCAT 1111
Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1112 CCACAGTTCACACTTTTGTCCCAAGAGATGTAGGGAAGACGGATATAAACGGCTAC 1171
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1172 ACTATTCTGTGAAGACCAAAAGTCATGTTAATGTTTGGGCATTTGGGAAGAGATCCGAAA 1231
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1232 TATTGGGATGATGCAGAAAGTTTTTAAGCCAGAGAGATTTGACACGGCTCTGTGCACCTT 1291
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1292 GTTGTAACAAATTTTGAGTATCTTCCCTTTGGCGGTGGGAGAAGATTGTCCCGGATA 1351
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1352 TCATTGGCTAGCAATGTTATTGTCGGTGTGGCTCATTTGTTATATCATCTTCGACTGG 1411
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1412 AAATCCCTATTGGAAATGGAGCCAAAGACTTGAACCTGACTGAATTTGGTTGGAGTAAT 1471
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1472 GCTGCCAGAAAAGATGACCTTATTTTGGTTGCCACTCCTCTATGAA 1516
```

```
RESULT 10
CS223950
LOCUS 1610 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 206 from Patent WO200511217.
ACCESSION CS223950
VERSION CS223950.1 GI:83685854
KEYWORDS
```

```
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
REFERENCE 1
AUTHORS Xu,D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 200511217-A 206 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
source 1..1610
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Alignment Scores:
Pred. No.: 6.51e-120 Length: 1610
Score: 1131.00 Matches: 222
Percent Similarity: 66.1% Conservative: 105
Best Local Similarity: 44.8% Mismatches: 154
Query Match: 43.9% Indels: 14
DB: 2 Gaps: 6
US-10-759-813-2 (1-500) x CS223950 (1-1610)
Qy 10 SerIleLeuIleSerPheLeuLeuValLeu---IleLeuValValMetArgLeuTrp 28
Db 37 TCTAACTTAGTTGCATCTCTCTCTCTCCAGCATCTTTCTTGTATTCAAAAAATGG 96
Qy 29 Lys-----LysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 97 AAAACCCAGAAATAAATTTGCCTCTGTCATGGAATAATTAACCTTTATTGGAAGTTTA 156
Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 157 CACCATTTGGCTGGCGAGGTCACCTTCTCCATATGGCCTTAAAAAATTTAGCCAAACGC 216
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 217 TATGGTCTCTTATGTCATTTACAACTTTCGACAAATTTCTACACTCGTCATATCATCACCT 276
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 277 CAATATGGCAAAAGAGTACTAAATACTCACACCTCGCTTTTGGCCACTAGACCAAGCTT 336
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 337 GTCGTGGCCGACATCATTCATCGACAGCAGGACATAGCACCTTTCGCCATACGGTGAA 396
Qy 127 HisTrpArgGlnMetLysIleIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 397 TACTGGAGACAAATTCGTAATAATTTGCATATTGGAACCTTTGAGTGCCCAAGATGTCGAA 456
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 457 TTTTATTAGCTCGATTCGCCAAGATGAGCTCGAAGATGGTTTTCATATACGAACAGCG 516
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 517 CCCAATCTTCAGTCATCTTACCAGCAAGATTTTGTGTTTACGAGTTCGGTAATTTGT 576
Qy 187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 577 AGATCAGCTTTTAGGGAAGATATGTGTCACCAAGACAAATTTATATGAGGGAA 636
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 637 ATAATATCATTTGGCAGGTGGATTTAGTATTGCTGATTTTTCCTACATGGAAATGATT 696
Qy 226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle 245
```

```
Db 697 CATGATATTGATGGTTCAAAATCTAAACTGGTGAAGGCACATCGTAAGATTGATGAAT 756
Qy 246 LeuGluGluLeuAsnGluHisLysAlaAsnLys----- 257
Db 757 TTGGAATAATGGTAATAGCACAACAGAAATCGAGCAGATGGTAAAAAGGGTAATGGT 816
Qy 258 ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnVal 277
Db 817 GAATTTGGTGGAGAAGATCTGATTGATGTTTGTAAAGATTAGACAAAAGTGCGAGAATT 876
Qy 278 ProValProValThrAsnGluSerLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db 877 CAAATTCCAATCAGATGACATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 936
Qy 298 SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
Db 937 TCGGAACATCATCGCACTATAATTTGGGCATTAGCTGAAATGATGAAGAAACCAAGT 996
Qy 318 GluLeuArgLysAlaGlnGluValArgGlnValPhe---GlyGluMetGlyLysVal 336
Db 997 GTTTAGCAAGGCACAAGCTGAAGTGAGCCAAAGCTTTGAAGGGGAAGAAATTAGTTT 1056
Qy 337 AspGluSerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArg 356
Db 1057 CAAGAGATTGATATTGATGAAGCTAAAGTATTTGAAGTTAGTCAAGAAACTTTAAGA 1116
Qy 357 LeuHisProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgLysAsp 375
Db 1117 ATGCACCTCCCAATTCCTCTGTTAGTCCCTAGAGAATGATGAAGATCAAGAAATGAT 1176
Qy 376 GlyTyrGluHisProAsnThrArgLysValValAsnAlaTrpAlaLeuGlyArgAsp 395
Db 1177 GGTTCAAATATACCTTTCAAAACAAGAGTCATTTGTAATGATGGCAATTCGACGAGAT 1236
Qy 396 ProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaLeu 415
Db 1237 CCTCAAGTGGGATGATCCTGAAAGCTTTACGCCAGAGAGATTTGAGAATAATTCATTT 1296
Qy 416 AspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgLysCysPro 435
Db 1297 GATTTCTTGGAAATCATCATCAATTTATTCATTTGGTGGAGGAAGAAAGATTTGTCTC 1356
Qy 436 GlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPhe 455
Db 1357 GGAATGCTATTTGGTTTACCTAATGTTGGACAACCTTTAGCTCAGTTACTTTATCACTTC 1416
Qy 456 AsnTrpGluAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaLeuGly 475
Db 1417 GATTGGAAACTCCCTAATGGACAAACTCACCAAAATTTTCGACATGACTGAGTCACCTGGA 1476
Qy 476 GlyAlaLeuArgLysLysIleAspLeuLysLeuIleProIlePro 490
Db 1477 ATTTCTGCTACAAAGAAAGATGATCTTATTTGATGGCACTCCT 1521

RESULT 11
LOCUS CS224014 1610 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 270 from Patent WO2005111217.
ACCESSION CS224014
VERSION CS224014.1 GI:83685886
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Solanales; Solanaceae; Nicotiana.
1
REFERENCE
AUTHORS Xu, D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 200511217-A 270 24-NOV-2005;
FEATURES
U.S. Smokeless Tobacco Company (US)
Location/Qualifiers
1. .1610
source
```

```
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN
Alignment Scores:
Pred. No.: 6,51e-120 Length: 1610
Score: 1131.00 Matches: 222
Percent Similarity: 66.1% Conservative: 105
Best Local Similarity: 44.8% Mismatches: 154
Query Match: 43.9% Indels: 14
DB: 2 Gaps: 6

US-10-759-813-2 (1-500) x CS224014 (1-1610)
Qy 10 SerIleLeuIleSerPheLeuLeuValLeu---IleLeuValValValMetArgLeuTrp 28
Db 37 TCTAACTTAGTTAGTTCATTTCTCTCTCTCCAGCATCTTTCTTCTTATTCAAAAATGG 96
Qy 29 Lys-----LysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 97 AAAACCCAGAAAATAAATTTGCTCTCTGTCTCATGGAAATTTACCTTTTATTGGAAGTTTA 156
Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 157 CACCATTGGCTGTGGCAGGTCCACTTCCTCACCATGGCTTAAAAAATTTAGCCAAACGC 216
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 217 TATGTCCTCTTATGATTTTCAACATTTGGACAAATTTCTTACACTTCGTATCATCACCT 276
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 277 CAAATGGCAAAAGAGTACTAAAACTCAGACCTCGCTTTTCCCACTAGACCAAGCTT 336
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 337 GTGCTGGCGGACATCATTCCTACGACGACGACGACATAGCACTTTTCGCCATACGCTGAA 396
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 397 TACTGGACAAATTTGTAATTTGCAATTTGGAATCTCTTGAGTGCCTCAAGTGTGCAAG 456
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 457 TTTTATTAGCTCGATTCGCCAAGATGAGCTCTCAAGATGGTTTCATCTATACGAACGACG 516
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 517 CCCAATCTTCCAGTCAATCTTACCACCAAGATTTTGTGTTTACGAGTTTCGGTAATTTGT 576
Qy 187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 577 AGATCAGCTTTAGGGAAGATATGTGTCGACCAAGCAAAATTCATCTTTTATGAGGGAA 636
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 637 ATAATATCATTCGCGAGGTGGATTAGTATTGCTGATTTTTCCTACATACGAAATGATT 696
Qy 226 HistyrIleIleGlyAlaGluSerLysProArgLysLeuHisGlnGluIleAspAspIle 245
Db 697 CATGATATTGATGGTTCAAAATCTAACTGGTGAAGGCACATCGTAAGATTGATGAAT 756
Qy 246 LeuGluGluIleLeuAsnGluHisLysAlaAsnLys----- 257
Db 757 TTGGAATAATGGTAAATAGCACAACAGAAATCGAGCAGATGGTAAAAAGGGTAATGGT 816
Qy 258 ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnVal 277
Db 817 GAATTTGGTGGAGAAGATCTGATTGATGTTTGTAAAGATTAGACAAAAGTGCGAGAATT 876
Qy 278 ProValProValThrAsnGluSerLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db 877 CAAATTCCAATCAGATGACATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT
```



```

Db      1061 CAGATTGATCTTGATGAGCTAAATATCTCAAGTTAGTATCAAGAAACCTTTAAGAATG 1120
Qy      358 HisProProVal--ValLeuileProAtGluCysArgGluThrThrArgileAspGly 376
Db      1121 CACCCTCCAAATCTCTATTAGTCTTAGAGAATGATGGAGGATCAAGAATGATGGT 1180
Qy      377 TyrGluileHisProAsnThrArgileValValAsnAlaTrpAlaileGlyArgAspPro 396
Db      1181 TACAATATACCTTTCAAAACAAGAGTCATAGTTAATGTCATGGCAATCGGACGAGATCCA 1240
Qy      397 AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaileAsp 416
Db      1241 GAAAGTTGGGATGACCCCGAAAGCTTTATGCCAGAGAGATTTGAGAGATTTCTATTGAC 1300
Qy      417 TyrLysGlyThrThrPheGluLeuValProPheGlyValaGlyLysArgileCysProGly 436
Db      1301 TTCTTTGGAAATCATCATGTTTATACCATTTGGTGCAGGAAGAAGATTTGTCGGGA 1360
Qy      437 IleThrSerAlaileThrAsnLeuGluTyrValileileAsnLeuLeuTyrHisPheAsn 456
Db      1361 ATGCTATTGGTTTAGCTAATGTTGGACAACCTTTAGCTCAGTTACTTTATCATCTTCGAT 1420
Qy      457 TrpGluLeuAlaAspGlyileThrProGlnThrLeuAspMetThrGluAlaileGlyGly 476
Db      1421 TGGAAACTCCCTAATCGACAAGTCTATGAGATTTTCGACATGACTGAGTCACCTGGAATT 1480
Qy      477 AlaLeuArgLysLysileAspLeuLysLeuileProileProTyr 491
Db      1481 TCTGCTACAGAAAGGATGATCTGTTTGTGATTCACCTCTTAT 1525

```

RESULT 13

```

LOCUS      CS223944
DEFINITION Sequence 200 from Patent WO2005111217.
ACCESSION  CS223944
VERSION     CS223944.1 GI:83685851
KEYWORDS
SOURCE      Nicotiana tabacum (common tobacco)
ORGANISM    Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1
AUTHORS     Xu,D.
TITLE       Nicotiana nucleic acid molecules and uses thereof
JOURNAL     Patent: WO 2005111217-A 200 24-NOV-2005;
            U.S. Smokeless Tobacco Company (US)
FEATURES    Location/Qualifiers
             1..1566
             /organism="Nicotiana tabacum"
             /mol_type="unassigned DNA"
             /db_xref="taxon:4097"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      1,22e-119      Length:      1566
Score:          1128.50      Matches:      218
Percent Similarity: 66.1%      Conservative: 109
Best Local Similarity: 44.0%      Mismatches: 155
Query Match:    43.8%      Indels:      13
DB:             2           Gaps:         5

US-10-759-813-2 (1-500) x CS223944 (1-1566)

Qy      10 SerileLeuileSerPheLeuValLeu---IleLeuValValValMetArgLeuTrp 28
Db      29 TCTAACTTAGTTCGATCTTGTCTTCTCTCCAGCATCTTCTTCTATTCAAAAATGG 88
Qy      29 Lys-----LysGlnAsnProProGlyProTrpLysPheProileileGlyAsnLeu 46
Db      89 AAACACGAAATAAATTGCTCTCTGTCATGGAATTAACCTTTTATGGAATTTA 148
Qy      47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnile 66

```

```

Db      149 CACCAITTTGGCTGTGCAGGTCACACTTCTCCACCATGGCTAAATAATTAGCAACGC 208
Qy      67 TyrGlyProValMetSerLeuGlnileGlyGlnValSerAlaValValleSerSerAla 86
Db      209 TATGGTCTCTTATGCAATTTCAAACTTGGACAAATTCCTACACTCATCATCATCACCT 268
Qy      87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProileVal 106
Db      269 CAAATGGCAAAAGAAAGTACTAAAACTCAGACCTCGCTTTTCCCACTAGCAACAGCTT 328
Qy      107 LeuAspAlaGlnileValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db      329 GTGTGGCCGCGCATCATCTACAGACAGCAGCAGCATAGCATTTTCTCGTACGGTAA 388
Qy      127 HisTrpArgGlnMetLysLysileTrpIleLeuGluPheLeuSerAlaLysValGln 146
Db      389 TACTGGAGACAAATTCGTAAAAATTTGCATATTTGGAACCTCTTGAGTCCCAAGATGTCAAA 448
Qy      147 SerSerArgLeuileArgGluGluMetGluAspAlaileThrPheLeuArgSerLys 166
Db      449 TTTTITTAGCTCGATTCGCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAACACA 508
Qy      167 AlaGlySerProValAsnileThrLysileileTyrGlyileileSerileMetIle 186
Db      509 CCAATCTTTACAGTCAATCTTACTGACAAAATTTTGTGTTTACGAGTTCGGTAACCTGT 568
Qy      187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db      569 AGATCAGCTTTTAGGGAAGATATGTGTGACCAAGACAAATATGATATTTTATGAGGGAA 628
Qy      206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db      629 ATAAATATCATTCGCGAGTGGATTTAGTATTTGCTGATTTTTCCTACATGGAAAAATGATT 688
Qy      226 HisTyrileileGlyAlaGluSerLysProArgArgLeuHisGlnGlnileAspAspile 245
Db      689 CATGATATTGATGGTTTGGAAATCTAAACTGGTGAAGACACATCGTAAGATTGATGAATT 748
Qy      246 LeuGluGluileLeuAsnGluHisLysAlaAsnLys----- 257
Db      749 TTGGGAAATTTGTTGTATGAGCACAAAAGAACAGACAGATGGCAAGAGGTAATGGT 808
Qy      258 ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnVal 277
Db      809 GAATTTGGTGGTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 868
Qy      278 ProValProValThrAsnGluSerileLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db      869 CAAATTCCTATCACAATGACAATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 928
Qy      298 SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
Db      929 TCTGAACATCATCGACGACTATAATTTGGGCAATTTAGCTGAATATGATGAAGAACCAAGT 988
Qy      318 GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
Db      989 GTTTTAGCAAGGCAACAGCTGAAGTAAAGCAAGCTTTTGAAGAGGAAAAAGGTTTTCAA 1048
Qy      338 GluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeu 357
Db      1049 CAGATTGATCTTGATGAGCTAAAAATATCTCAAGTTAGTAAATCAAGAAAAACCTTAAGAATG 1108
Qy      358 HisProProVal---ValLeuileProArgGluCysArgGluThrThrArgileAspGly 376
Db      1109 CACCCTCCAAATTCCTCTATTAGTTCTTAGAGAATGATGATGGAGGATACAAAGATTTGATGT 1168
Qy      377 TyrGluileHisProAsnThrArgileValValAsnAlaTrpAlaileGlyArgAspPro 396
Db      1169 TACAATATACCTTTCAAAACAAGAGTCATAGTTAATGATGGCAATCGACAGATCCA 1228
Qy      397 AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaileAsp 416

```


Search completed: May 31, 2006, 09:03:53
Job time : 8518 secs

This Page Blank (uspto)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: May 31, 2006, 06:39:28 ; Search time 6894 Seconds
(without alignments)
6083.478 Million cell updates/sec

Title: US-10-759-813-2
Perfect score: 2574
Sequence: 1 MEQKNLSPSILFLVLI.....KIDKLPIPYQVSLGSNIS 500

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=abss/ABSSWEB spool/US10759813/runat_31052006_060306_2636/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US10759813 @CN 1.1 6323 @runat_31052006_060306_2636 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: gb_est1: *
2: gb_est3: *
3: gb_est4: *
4: gb_est5: *
5: gb_est6: *
6: gb_hic: *
7: gb_est2: *
8: gb_est7: *
9: gb_est8: *
10: gb_est9: *
11: gb_gss1: *
12: gb_gss2: *
13: gb_gss3: *
14: gb_gss4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	954.5	37.1	1548	13	CL965804
2	947.5	36.8	1569	13	CL965067
3	947	36.8	1833	6	AY104522
4	934	36.3	1560	13	CL965799

5	927.5	36.0	1536	13	CL965911	CL965911	OsIFCC013
6	927	36.0	1575	13	CL977750	CL977750	OsIFCC030
7	923.5	35.9	1542	13	CL966810	CL966810	OsIFCC014
8	913.5	35.5	1557	13	CL966924	CL966924	OsIFCC014
9	910	35.4	1536	13	CL962214	CL962214	OsIFCC007
10	909.5	35.3	1581	13	CL967011	CL967011	OsIFCC014
11	907	35.2	1530	13	CL965809	CL965809	OsIFCC012
12	905	35.2	2259	13	CL965071	CL965071	OsIFCC011
13	901	35.0	1518	13	CL965068	CL965068	OsIFCC011
14	898	34.9	1539	13	CL965064	CL965064	OsIFCC011
15	897	34.8	1557	13	CL958628	CL958628	OsIFCC001
16	895	34.8	1572	13	CL965801	CL965801	OsIFCC012
17	890.5	34.6	1527	13	CL964411	CL964411	OsIFCC010
18	890.5	34.6	1536	13	CL968663	CL968663	OsIFCC017
19	888	34.5	1542	13	CL981555	CL981555	OsIFSC045
20	885	34.4	1593	13	CL960539	CL960539	OsIFCC004
21	881.5	34.2	1545	13	CL965805	CL965805	OsIFCC012
22	871	33.8	1539	13	CL958604	CL958604	OsIFCC001
23	856	33.3	1572	13	CL977749	CL977749	OsIFCC030
24	855	33.2	1572	13	CL965811	CL965811	OsIFCC012
25	849	33.0	1518	13	CL977242	CL977242	OsIFCC030
26	843.5	32.8	1560	13	CL965062	CL965062	OsIFCC011
27	841.5	32.7	1566	13	CL947270	CL947270	OsIFSB003
28	841.5	32.7	1566	13	CL965806	CL965806	OsIFCC012
29	839.5	32.6	1629	6	CNSOABBN	CL965806	Arabidops
30	839	32.6	1902	13	CL965800	CL965800	OsIFCC012
31	836	32.5	732	10	DV153215	DV153215	CV03091A2
32	831.5	32.3	1518	13	CL964423	CL964423	OsIFCC010
33	830	32.2	1362	13	CL965069	CL965069	OsIFCC011
34	814	31.6	1614	13	CL958110	CL958110	OsIFCC000
35	812.5	31.6	1422	13	CL965810	CL965810	OsIFCC012
36	811	31.5	1569	13	CL979405	CL979405	OsIFCC033
37	807.5	31.4	1656	6	CNSOAG90T	CL965800	Arabidops
38	803.5	31.2	1565	6	CNSOAG85	CL965800	Arabidops
39	803	31.2	1580	6	CNSOABYN	CL965800	Arabidops
40	800.5	31.1	1643	6	CNSOAB1J	CL965070	OsIFCC011
41	800	31.1	1479	13	CL965070	CL965070	OsIFCC011
42	796.5	30.9	1532	6	CNSOAG1J	CL965070	OsIFCC011
43	795	30.9	1602	13	CL958109	CL958109	OsIFCC000
44	785	30.5	1401	13	CL965803	CL965803	OsIFCC012
45	782	30.4	1584	13	CL963214	CL963214	OsIFCC008

ALIGNMENTS

RESULT 1	CL965804	1548 bp	DNA	linear	GSS 21-SEP-2004
LOCUS	OsIFCC012776	Oryza sativa Express Library Oryza sativa (indica cultivar-group)	genomic, genomic survey sequence.		
DEFINITION	CL965804	GSS.			
ACCESSION	CL965804.1	GI:52386289			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					

```

FEATURES
  source
    Class: exon-trapped.
    Location/Qualifiers
      1..1548
        /organism="Oryza sativa (indica cultivar-group)"
        /mol_type="genomic DNA"
        /db_xref="caxon:39946"
        /clone_lib="Oryza sativa Express Library"
        /note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.:      1,558-98      Length:      1548
Score:          954.50      Matches:      198
Percent Similarity: 61.1%      Conservative: 105
Best Local Similarity: 39.9%      Mismatches: 170
Query Match:      37.1%      Indels:      23
DB:              13          Gaps:       7

US-10-759-813-2 (1-500) x CL965804 (1-1548)
Qy 15 PheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGlnAsnPro--- 33
Db 15 PheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGlnAsnPro--- 33
Qy 43 TTCCTCTCCTGCTTATGTTAGTTAGGACGACGAGCGCGTGGCGACATAAT 102
Db 43 TTCCTCTCCTGCTTATGTTAGTTAGGACGACGAGCGCGTGGCGACATAAT 102
Qy 34 -----ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHis 48
Db 34 -----ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHis 48
Qy 103 GCGCGGTGAAGTTGCGCGTGGCGGTGGCGACTGCGCGCTGTCGCGCAACCTCCACCAG 162
Db 103 GCGCGGTGAAGTTGCGCGTGGCGGTGGCGACTGCGCGCTGTCGCGCAACCTCCACCAG 162
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuA---GlnIleTyr 67
Db 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuA---GlnIleTyr 67
Qy 163 GTACATGCGCGCGCGCGCTGTCGACCGCACCATGATGCGCGCGCGCGCTCGAC 222
Db 163 GTACATGCGCGCGCGCGCTGTCGACCGCACCATGATGCGCGCGCGCGCTCGAC 222
Qy 68 GlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGlu 87
Db 68 GlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGlu 87
Qy 223 GCGCGCTCATGTCGCTCAGGCTCGCGAGGTCGCCGTCGTCGCTGTCGTCGCGCGAC 282
Db 223 GCGCGCTCATGTCGCTCAGGCTCGCGAGGTCGCCGTCGTCGCTGTCGTCGCGCGAC 282
Qy 88 AlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeu 107
Db 88 AlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeu 107
Qy 283 GCGCGCAGGAGATCACCAGACGACGACGTCGCGCTTCGCGACACGCGCGTGGAGCCCG 342
Db 283 GCGCGCAGGAGATCACCAGACGACGACGTCGCGCTTCGCGACACGCGCGTGGAGCCCG 342
Qy 108 AspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHis 127
Db 108 AspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHis 127
Qy 343 ACCAGCGCGCGCTCGGTGCGAGCGGCGGCGGTGTTGCGACGACGACGCGCGGTG 402
Db 343 ACCAGCGCGCGCTCGGTGCGAGCGGCGGCGGTGTTGCGACGACGACGCGCGGTG 402
Qy 128 TrpArgGlnMetLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 147
Db 128 TrpArgGlnMetLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 147
Qy 403 TGGCGCCAGCTCCGCAAGCTGTCGTCGAGCTGCTCGGCGCGCGCGCGTCCGCTCG 462
Db 403 TGGCGCCAGCTCCGCAAGCTGTCGTCGAGCTGCTCGGCGCGCGCGCGTCCGCTCG 462
Qy 148 SerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu---ArgSerLys 166
Db 148 SerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu---ArgSerLys 166
Qy 463 TTCCGCGCGCTCGCGGAGGAGCGCGCGCTCGTCGCGCGCGCTCGCGCGTCCCGG 522
Db 463 TTCCGCGCGCTCGCGGAGGAGCGCGCGCTCGTCGCGCGCGCTCGCGCGTCCCGG 522
Qy 167 AlaGlySerProValAsnIleThrLysLysLysLysLysLysLysLysLysLysLys 186
Db 167 AlaGlySerProValAsnIleThrLysLysLysLysLysLysLysLysLysLysLys 186
Qy 523 CCGCGCAGGCGCTGAACCTACGACGCGGATACGCGCGGTGATCAGGACGCGAGTG 582
Db 523 CCGCGCAGGCGCTGAACCTACGACGCGGATACGCGCGGTGATCAGGACGCGAGTG 582
Qy 187 ArgThrSerValGlyAsn-----CysLysGlnLysGluArgLeuLeuSerValAlaAsp 204
Db 187 ArgThrSerValGlyAsn-----CysLysGlnLysGluArgLeuLeuSerValAlaAsp 204
Qy 583 CCGCGCATGATCGGGGACAGGTTCCGGCGCGCGCGAGTTCCTGGAGTGTCTCGCCGAC 642
Db 583 CCGCGCATGATCGGGGACAGGTTCCGGCGCGCGCGAGTTCCTGGAGTGTCTCGCCGAC 642
Qy 205 AlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeu 224
Db 205 AlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeu 224
Qy 643 ATCGTC---AAGATCGCGTCCCGGTTTCAGCTTCGACGACCTGTTCCCGTGGGGTCTC 699
Db 643 ATCGTC---AAGATCGCGTCCCGGTTTCAGCTTCGACGACCTGTTCCCGTGGGGTCTC 699
Qy 225 LeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnIleLeuAspAsp 244
Db 225 LeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnIleLeuAspAsp 244
Qy 700 GCGCGCGCATCGCGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 759
Db 700 GCGCGCGCATCGCGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 759
Qy 245 IleLeuGluGluLeuAsnGluHisLysAlaAsnLysProPheGluAla----- 261
Db 245 IleLeuGluGluLeuAsnGluHisLysAlaAsnLysProPheGluAla----- 261
Qy 760 CTGATGACAGCGGTGTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 819
Db 760 CTGATGACAGCGGTGTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 819
Qy 262 -----AspAsnLeuMetAspValLeuLeuAsnLeuGlnLys 273
Db 262 -----AspAsnLeuMetAspValLeuLeuAsnLeuGlnLys 273

```

```

Db 820 GACGGCGCCATGGACGACGCGAGGAGGACTCTGTGACGTGCTCTTCAGGATACAGAA 879
Qy 274 AsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMet 293
Db 880 GATGGTGGCTCGAGGTGCTCTCACCATTGGCAACATCAAGCAATCCTCTGGACCTC 939
Qy 294 PheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMet 313
Db 940 TTCAATGCAGGAGTGAGACATCAGCAATACGCTCCAGTGGGTATGTCTCAGAGCTAATG 999
Qy 314 LysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMet 333
Db 1000 AGGAACCCAAAGATGATGAGAAAACCAAGACGAGTTAAGGAACAACCTACAGGGGAG 1059
Qy 334 GlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGlu 353
Db 1060 ACAACGGTTAACCGAAGACGATCTCACAATCTAAGTACCTGAAGCTTGTCTATCAAGGAG 1119
Qy 354 ThrLeuArgLeuHisProProVal---ValLeuIleProArgGluCysArgGluThrThr 372
Db 1120 ACGCTTAGGTTACATCCGCTGCTGCTTACTTCTCCAGAGAGTGCACGAAGCATGC 1179
Qy 373 ArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIle 392
Db 1180 AATGTATCGGGTACGACGCTCCCAATACACTACCGTGTATTAACGTTGGGCAATC 1239
Qy 393 GlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAsp 412
Db 1240 AACAGGACCCCAATATTGGACATGGCCGAGATGTTTAAAGCCAGAACGCTTTCACAAT 1299
Qy 413 CysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyValGlyLysArg 432
Db 1300 AGCATGATTGATTCAAGGTACAGACTTCGAGTTCGTCGCAATTTGGGGCTGGAGAAGG 1359
Qy 433 IleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeu 452
Db 1360 ATCTGCTCGCATAGCATTTGCACAGTCCCAATATGAGCTTGTGCTGTACATCTCTA 1419
Qy 453 TyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGlu 472
Db 1420 TACCACITTAGCTGGAGCTTCTTAGTGTATGTCCTGAGAGCTTAGATATGACTGAG 1479
Qy 473 AlaIleGlyAlaLeuArgLysLysLysLysLysLysLysLysLysLysLysLysLys 488
Db 1480 GACATGGGACTCTCTCGGAAGGAAGAAATGATCTTTACCTACATCCCA 1527

RESULT 2
LOCUS CL965067 1569 bp DNA linear GSS 21-SEP-2004
DEFINITION OSIFCC011616 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL965067
VERSION CL965067.1 GI:52384824
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BBP clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 1569)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676

```


JOURNAL PUBLISHED
REFERENCE AUTHORS
TITLE
JOURNAL REFERENCE
AUTHORS TITLE
JOURNAL COMMENT
FEATURES source

Plant Physiol. 134 (4), 1317-1326 (2004)
15020742
2 (bases 1 to 1833)
Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
3 (bases 1 to 1833)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

Location/Qualifiers
1..1833
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeGDB:634293"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Corneausus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Alignment Scores:
Pred. No.: 1.48e-97 Length: 1833
Score: 947.00 Matches: 195
Percent Similarity: 61.3% Conservative: 112
Best Local Similarity: 38.9% Mismatches: 168
Query Match: 36.8% Indels: 26
DB: Gaps: 12

US-10-759-813-2 (1-500) x AY104522 (1-1833)

Qy 11 IleLeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLys 30
Db 29 GTCTACTCGGACTGGCGCTGTGTGCTGCTGCTTGTGTGTCGCGNAG---CGCGGG 85
Qy 31 GlnAenPro-----ProProGlyProTrpLysPheProIle 42
Db 86 CGCGGCCGGTGGAACGGCCACGCCTCGCGCTCGCGCGGGCGGTGGCACGTCCCAATC 145.

Qy 43 IlegAenLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAla 62
Db 146 ATCGCGACATGACACCAT-----CTCGCGGCCAGCTCTCCACC GGCGATCGCGAC 199

Qy 63 LeuAlaGlnIleTrgYlryProValMetSerLeuGlnIleGlyGlnValSerAlaVal 82
Db 200 CTGGCGGGCGCCACGGGGCGGTGATGCTGCTCGCGTCGCGAGGTGCGCAGCTGGT 259

Qy 83 IleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGln 102
Db 260 GTCTCTCTCCCGAGGCGCGCGGAGGTGATGAAGACGCACACCGCGGTTCGCGTCC 319

Qy 103 ArgProIleValLeuAspAlaGlnIleValPheTyrsenAsnArgLysAspValLeuPheAla 122
Db 320 CCGCGCTGACGGCCACCGTGGCGGTGTGACGAACAACGCCCGGGACATCATCTTCGG 379

Qy 123 SerTyrglyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAla 142
Db 380 CCCTACGGGGAGACTGTCGCGCAGCTGCGGAAGTTGGCCATCACGGAGCTCTGAGCGG 439

Qy	143	LysLysValGlnSerSerArgLeuIleAtrgGluGluMetGluaspAlaIleThrPhe	162
Db	440	CGCGGGTGTCTCTTTCGCGCATCCGCGAGGAGGTTCGGGCCACGCTGCGCGCG	499
Qy	163	LeuArgSerLysAlaGlySer-----ProValAsnIleThrLysIleIleTyr	178
Db	500	TGCGCGCGCCGCGCGGAGTGC CGCGCCCGTGGAGATGCGCGCGCGTGTCC	559
Qy	179	GlyIleIleIleSerIleMetIleAtrgSerValGlyAsn---CysLysGlnLysGlu	197
Db	560	CGCGTCTGTGGCCAGCGCACGCGTGC CGCGCTGTATGGCGCACCGGTGCAGGGAC	619
Qy	198	ArgLeuSerValAlaaspAlaValAsnGlnLualaAalaThrSerPheGlyThrAlaAsp	217
Db	620	GTGTTCCTCGGAGCTCGACCGCTCATCGGGCTATCGCGGGGTTC AACCCGCGCGAC	679
Qy	218	AlaPheProThrTrpLysLeuHisTyrIleIleGlyAlaGluSerLysProAtrgArg	237
Db	680	CTGTGGCGCTCTCCGCTCTGTGGCGCTGCTGAGCGCGCGCTGCGCGCGCGAGGAG	739
Qy	238	LeuHisGlnGluIleAepAspIleLeuGluGluIleLeuAsnGluHis-----LysAla	255
Db	740	TGCGGGACACGCTGTTCGGGATCTCGACGGCATCATCAACGAGCACCTGCGAGAGATG	799
Qy	256	AsnLysProPheGluAla-----AspAsnLeuMetAspValIleLeuAsnLeuGlnLys	273
Db	800	GACAGCGCGCGGAGGCTGCAGCGAGGACCTTCTCGACGTGCTGCTCAAGATACACACAG	859
Qy	274	AsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMet	293
Db	860	GATGGCAGCCTCCAAATCCGCTCGACATGGACGTGCTTAAGCGCTCATCTTTGACATC	919
Qy	294	PheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMet	313
Db	920	TTGGCGCGCGCAGCAGACGTGCGGCACACACGTGGAGTGGGCCATGCGCGGAGCTGATC	979
Qy	314	LysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMet	333
Db	980	CGGAACCCCAAGCGATGCAGCGGGCGACCGCGAGGTGCGCAGGCTTCGGCGCGCGC	1039
Qy	334	GlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGlu	353
Db	1040	GGCGCGTGGCGAGCAGAGCTGGCGAGCTGCGGTACCTGCATCTCGTGATCCGGGAG	1099
Qy	354	ThrLeuArgLeuHisProProVal---ValLeuIleProArgGluCysAtrgGluThrThr	372
Db	1100	ACGTTCGGGTGCACCGCGCGCTGCGCGTCTGCTCCCGCGGAGTCCCGAGGACCGCGC	1159
Qy	373	ArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIle	392
Db	1160	CGCGTGTGGGTACGACGTGCCCGGGCGACACCGTGTGTGTGAACGTCTGGCGCGTG	1219
Qy	393	GlyArgAspProAsnThrTrp---SerGluProGlyLysPheAsnProGluAtrgPhe---	410
Db	1220	GGCGCGCAGCGAGGTACTGGCGGGCGACCCCGAGGAGTTCGGCGCGCGAGCGGTTCGAG	1279
Qy	411	---LysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyVala	429
Db	1280	GCCGAGCCAGCGGGTGGAGTTCAAGGGCGCGACACTTCGAGTCTCTGCGGTTCGGCGCC	1339
Qy	430	GlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIle	449
Db	1340	GGCCGGAGGATGTCCCGGGCATGGCGTTCGGGCTCCCAACGTGCGAGCTCGCGCTGCC	1399
Qy	450	AsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThr---ProGlnThrLeu	468
Db	1400	AGCTGTCTGCCATCTCGACTGGGAG---GGCGCGCGGTGTCTGACCCCGCGCGAGTTC	1456
Qy	469	AspMetThrGluAlaIleGlyGlyAlaLeuArgLysLysIleAspLeuLysLeuLeuPro	488
Db	1457	GACATGCCGAGGGGTTCATCACC CGCAGAGCGGAAGCGCACCTCTCTTTCGGCCCC	1516


```

Db      1393  GAACTCGCCCTCGCCACCCTCTTACCACCTTCGATTGGAGCCTCCCTGCTGGGGTGAAG 1452
Qy      465  ProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLysLysIleAspLeu 484
Db      1453  CCAAGTAGCTAGATATACCCAGAGGATGGGCATCACTGTTAGGAGGAGNAGATGACCTA 1512
Qy      485  LysLeu 486
Db      1513  TACCTA 1518

RESULT 5
CL965911
LOCUS   CL965911                1536 bp      DNA      linear      GSS 21-SEP-2004
DEFINITION  OsIFCC013000 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION CL965911
VERSION   CL965911.1  GI:52386501
KEYWORDS  GSS.
SOURCE    Oryza sativa (indica cultivar-group)
ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 1536)
AUTHORS   Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
JOURNAL   Unpublished (2004)
COMMENT   Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488876
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES  Location/Qualifiers
            source          1..1536
                        /organism="Oryza sativa (indica cultivar-group)"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:39946"
                        /clone_lib="Oryza sativa Express Library"
                        /note="Oryza sativa exon trapped genomic sequences "
ORIGIN
Alignment Scores:
Pred. No.:      2,036-95      Length:      1536
Score:          927.50      Matches:    199
Percent Similarity: 58.3%      Conservative: 100
Best Local Similarity: 38.8%      Mismatches:  181
Query Match:    36.0%      Indels:     33
Db:             13      Gaps:      12

US-10-759-813-2 (1-500) x CL965911 (1-1536)
Qy      6  LeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuLeu----- 21
Db      10  CTCCTTGACCAAGCGCTCTCTCTCCGGCTCGCGCTCGCGGTGCTTCAGATCGTCAAG 69
Qy      22  ValValMetArgLeuTrpLysLysGln-----AsnProProProGly 36
Db      70  CTCGCCGTCGTCAACAGAGGGAAGCAAGCGCGCGGTGTCACCGACGCCCGCGGG 129
Qy      37  ProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGly 56
Db      130  CCATGGAGGCTCGCGGTGATCGGACATGCAACAC-----CTCGCCGGGAAGTCGGC 183
Qy      57  HisGluArgPheArgAlaLeuAlaGlnIleTrpGlyProValMetSerLeuGlnIleGly 76
Db      184  CACC CGCGCTGAGGAGCCTCGCCCGCTGCGACCGCCCGCTGATGCTCCAGCTCGGC 243

```

```

Qy      77  GlnValSerAlaValValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGln 96
Db      244  GAGACGCGCTGGTGGTGGTTCAGGAGGAGTGGCGCGGAGGTGCTCAGGAGCGAC 303
Qy      97  AlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTrpAsnArg 116
Db      304  GACGCCAACTTCGCCACCCGCCCAAGGCTGTCGCGCGCGAGGTGCTCTTACCGCGGC 363
Qy      117  LysAspValLeuPheAlaSerTyxGlyAspHisTrpArgGlnMetLysLysIleTrpIle 136
Db      364  GCCGACATCTCTTCTCGCGCTCGCGGAGTACTGGCGCAAGCTCCGCGAGCTCTCGCC 423
Qy      137  LeuGluPheLeuSerAlaLysValGlnSerSerArgLeuIleArgGluGluGluMet 156
Db      424  GCCGAGGTGCTGGGCGCGAAGCGGTCTCTCTTCCGCGCATCAGGAAACAGAGATG 483
Qy      157  GluAspAlaIleThrPheLeuArgSerLysAlaGlySer---ProValAsnIleThrLys 175
Db      484  GCGAGTCGGGTGGAAAGGATCCGTGCGGTCCGGACCATCGGTCCCGTGATGTTAGTGA 543
Qy      176  IleIleTrpGlyIleIleIleSerIleMetIleArgThrSerValGlyAsnCysLysGln 195
Db      544  CTTTCTATGACATGGCAATCAGTATCGTCTCATGTGCATCTCTTTGGTAAG---AAACA 600
Qy      196  Lys-----GluArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPhe 213
Db      601  AGGAATGTGACGAGTACCTTTTCAGCAATAAAGACTGGGATCAGTCTGGTAGTGGTGT 660
Qy      214  GlyThrAlaAspAlaPheProThrTrpLys---LeuLeuHisTyxIleIleGlyAlaGlu 232
Db      661  AAGATCCCTGACCTCTTCCCGACATGGAAACGGTGTCTGTGCGAGTCACGGGCTATGG 720
Qy      233  SerLysProArgArgLeuHisGlnGluIleAspIleLeuGluGluIleLeuAsnGln 252
Db      721  CGTCCCTTGAANAATGTCATAGGATAGTGACCTCCACCTTGGAAAGAGGTCTACAGGAG 780
Qy      253  HisLysAla-----AsnLysProPheGluAlaAspAsn 263
Db      781  AGGAGAGCGCTGCAAGAGAAATGCAAGGCGCAGATTGGACATGGAGGACAAACGAGGAAA 840
Qy      264  LeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValThrAsn 283
Db      841  TTGTGCGATGTCATCAAAATTCATGAGCAAGGTGGC-----CACCTAAGTAGA 891
Qy      284  GluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSerLys 303
Db      892  AATAGCATCAAGTCGGTGATTTGATATGTTTCACAGCCGGAACAGGCACGTTAGCATCA 951
Qy      304  AlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGln 323
Db      952  TCGCTAAATTTGGGTATGTCGAACTAATAGGAACCCCAAGGGTGTACCAAGCTGCA 1011
Qy      324  GluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSer-----ArgPhe 341
Db      1012  GGTGAGATCCGAGAGCGCTTTCACGGCAAGGCTACCGTAGGGAGGGTGACATCCAGGTG 1071
Qy      342  HisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisProVal 361
Db      1072  AGCAACCTCTCTTACCTTAGGCTCTTTATCAAGGAGACCTTCGGTGTGCCCTCCAGTG 1131
Qy      362  ---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGluLeuHis 380
Db      1132  CCTCTCTTGTGCCCCGGGAGGACCATGTCATGTGCGAGGTCAATGGGTATACGATCCCT 1191
Qy      381  ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db      1192  GCAAGGTCACTATCTGTTAAATGCATGGGCTATTGGGAGAGACCCAAAGTATTGGGAT 1251
Qy      401  GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLysGlyThr 420
Db      1252  GATCCTGAGGAATTCAGGCCAGAACGATTTGAGGGCAACAAGGTGGACTTTTCAGGACT 1311

```

```

QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgCysProGlyLeuThrSerAla 440
Db 1312 AGCTACGAGTACCTACCTTTTGGAGCGGACGAAGGATTGCGCGGTATTACTACGCC 1371
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1372 TTACCTGTCTTGAGATACACTTGTGCAGCTTATATACCACTTCAATTGGTCACTCCCA 1431
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1432 AAGGGTGTACCC-----GAGGTAGACATGGAGGAGCACTTGGCTTGGTCACGCGT 1485
QY 481 LysIleAspLeuLysLeuLeuProIleProTyrGlnVal 493
Db 1486 ATGACGCGCCCTGCTACTATGTGCCACTCCATTGTTGTT 1524

RESULT 6
CL977750
LOCUS CL977750 1575 bp DNA linear GSS 21-SEP-2004
DEFINITION OsIFCC030951 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL977750
VERSION CL977750.1 GI:52409999
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BFP
Clade; Ehrhartoideae; Oryzae; Oryza.
1 (bases 1 to 1575)
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
Location/Qualifiers
source 1..1575
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.: 2,4e-95 Length: 1575
Score: 927.00 Matches: 204
Percent Similarity: 59.6% Conservatives: 103
Best Local Similarity: 39.6% Mismatches: 178
Query Match: 36.0% Indels: 30
DB: 13 Gaps: 11

US-10-759-813-2 (1-500) x CL977750 (1-1575)

QY 1 MetGluGlnLysAsn---LeuSerPheProSerIleLeuIleSerPheLeuLeuValLeu 19
Db 1 ATGGAGGACAAACGCGCGCTGCTGACCGTACTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 51
QY 20 IleLeuValValValMetArgLeuTrpLys-----LysGlnAsnProPro 34
Db 52 ATCGCGGTGCTCATCTCCAGCTGGTATCATATATGACCAACCAAGGCTCAACTGCCG 111
QY 35 ProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuThrSerAsp 54

```

```

Db 112 CCGGGGCCATGAAGCTGCGGTGATCGCGAGCTCCACCAC---CTGTGGGGAGCCAC 168
QY 55 LeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGly-----ProValMetSer 72
Db 169 GCCATCCACCGGTGCGGTGCGCGCTGCGCGAGAACGACGCGGCGCCACCACTTCATGCG 228
QY 73 LeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAlaAlaLysGluVal 92
Db 229 ATATCGCTCGGAGGTGTTCCCGGTGCTGCTGCTGCGGAGCGCGGCGGAGAGATC 288
QY 93 MetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleVal 112
Db 289 CTCAGGAACAGGACGTCACGTTGCGCGACGCTTCTCTCAGCACCACTTCGGGTGATC 348
QY 113 PheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLys 132
Db 349 ACCTTTCGGCGGCAACACATGCGCTTTCGCGCCCTACGCGGAGCGGTGGCGGAGCTCCG 408
QY 133 LysIleTrpIleLeuGluPheLeuSerAlaLysValGlnSerSerArgLeuIleArg 152
Db 409 AAGCTGTGCACGCTGGAGCTCTGTCCGCGCGGCGGTGAGTCTGTCGCGGCGATCCCG 468
QY 153 GluGluGluMetGluAspAlaIleThrPheLeuArgSerLys-----AlaGlySerPro 170
Db 469 GAGGAGGAGTGGCGCGCTGCTGCGGGACCTCGCGCGCTCGCGCGCGCGCGGAGGCC 528
QY 171 ValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIleArgThrSerVal 190
Db 529 GTGAACCTGAGCGGCGAGGATCGAAGCTTCATCAACGAGCTGCTGCTGAGGTCTCGCTC 588
QY 191 ---GlyAsnCysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsnGluAla 209
Db 589 GCGCGCGGAGGAGCACCGCGACGAGTCTCTCGACCGCTCGCGACCGCGCTTGGACCG 648
QY 210 AlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHisTyrIleIle 229
Db 649 ACGACATGGCTCACCGTGGCGGACGCTTCTCCGTGAGCAAGCTCGCGCGGATGCTCGCG 708
QY 230 GlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIleLeuGluIle 249
Db 709 ACGGCGCGCGCAAGCGCTCGCCAGCGGAGAGATAGACACATCTTGGAGCAGATC 768
QY 250 LeuAsnGlu-----HisLysAlaAsnLysProPheGlu 260
Db 769 ATTCAAGAACGAAGAGGATCATGACAGGACGAGCTTGGCGGCGCGCGCGCGGAG 828
QY 261 AlaAspAsn-----LeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsn 276
Db 829 GCGATGAACACGAGCGAGTGTCTTGGATGTCTGCTGAGGCTTCAGAAAGGATGCGAAC 888
QY 277 ValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAla 296
Db 889 ACACCAATCCCAATTACTAACGAAGTCATGCTGCTTTTATTATGACATGTTTTCCAGG 948
QY 297 GlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnPro 316
Db 949 GGCACGAGACATCATCTTCCACATTAATCTGACANTGGCAGAACTCATCAAGAAACCA 1008
QY 317 ThrGluLeuArgLysAlaGlnGluValArgGlnValPhe---GlyGluMetGlyLys 335
Db 1009 AAAGTAATGGCAAGGCTCATGTTCAAGTGGCGCAAGCTTTCAGAGAAAGAACCATC 1068
QY 336 ValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeu 355
Db 1069 ACAGAGGATGACGGTGTAAACGAGTTGACCTATCTCAAAATGTTGATTAAGGATCATTTG 1128
QY 356 ArgLeuHisProProValValLeuIle---ProArgGluCysArgGluThrThrArgIle 374
Db 1129 AGGATGCACTGCCAGTGCCTTTTAGCCCTCTGTAATGTGAGAGACATGCAAGTT 1188
QY 375 AspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArg 394

```

```

Db      1189  ATGGGCTACGACATACCGAAGACACAACTGTATTTTGTGAATGCGATGGCGCAATATGCAGA 1248
Qy      AspProAenThrTrpSerGluProGlyLysAenProGluArgPheLysAspCysAla 414
Db      1249  GACCCAAATATTGGATGATGCCGAAGATTCAACCCAGAGAGTTTGAGNACAATAAGT 1308
Qy      IleAspTyrLysGlyThrPheGluLeuValProPheGlyAlaGlyLysArgIleCys 434
Db      1309  ATCGACTATAAGGAAGTAACCTTTGAATTCCTTCATTCCTTGGATCGGCGGTAGATGTGT 1368
Qy      ProGlyIleThrSerAlaIleThrAenLeuGluTyrValIleIleAenLeuLeuTyrHis 454
Db      1369  GCGGCTATGAACCTTGGCATTTGCGAATGTAGAGCTCCCTTAGCTAGTACCTTCTGTATCAT 1428
Qy      PheAenTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIle 474
Db      1429  TTTGATTGGAAGTTGCTGATGGAATGATGCTGAAGATGTTGACATCAGGATGCACCT 1488
Qy      GlyGlyAlaLeuArgLysLysIleAspLeuLysLeuIleProIle 489
Db      1489  GGAATACTTGTAGTAAACGCAATAGCCTAATATATGTGCCCGGTT 1533

```

RESULT 7

```

CL966810      1542 bp      DNA      linear      GSS 21-SEP-2004
LOCUS      OsIFCC014149 Oryza sativa Express Library Oryza sativa (indica
DEFINITION      cultivar-group) genomic, genomic survey sequence.
ACCESSION      CL966810
VERSION      CL966810.1 GI:52388272
KEYWORDS      GSS.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM      Oryza sativa (indica cultivar-group)

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1542)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488876
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers

```

```

FEATURES
source
1..1542
/organism="Oryza sativa (indica cultivar-group)"
/mo_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      5,918-95      Length:      1542
Score:      923.50      Matches:      191
Percent Similarity:      57.3%      Conservative:      99
Best Local Similarity:      37.7%      Mismatches:      197
Query Match:      35.9%      Indels:      19
Db:      13      Gaps:      6

```

```

US-10-759-813-2 (1-500) x CL966810 (1-1542)

```

```

Qy      1 MetGluGlnLysAenLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeu 20
Db      1 ATGGAGGACACACGATACCTACTACTA---TCTCTGGCCTTGTCCATGCTCTCCGCCATCTCTC 57

```

```

Qy      21  LeuValValMetArgLeuTrpLysLysGln-----AenProProProGlyProTrp 38
Db      58  CTCTCCAAGCTGGTCTCGATTCCAAAGAACCCAGGCTGAACCTTCCCGCGGCGCATGG 117
Qy      39  LysPheProIleIleGlyAenLeuProHisIleLeuLeuThrSerAspLeuGlyHisGlu 58
Db      118  ACGTTCGCGGTGATCGGCAGCATCCACATCTCGCCAGCAACCCGAAACACC---CACAGG 174
Qy      59  ArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnVal 78
Db      175  GCGCTGCGGCTCTCTCGCAGAACACACGCGGCCCTGATGCAGCTGTGCTGGCGAGGTG 234
Qy      79  SerAlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAsp 98
Db      235  CCCCGGTGGTGGGTTCGACCGCGAGCGCGCGGAGATCCTCAGAAACACGAGACCTC 294
Qy      99  AlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAsp 118
Db      295  AGGTTTCGCGGACCCGACGTCACAGCACCGTCGCGAGCGTCTCTTCGACGCCAGCGAC 354
Qy      119  ValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGlu 138
Db      355  ATCTTCTTCTCCCGTACGCGAGCGGTGGCGCAGCTGCGCAAGCTGTGCACGCGAGAG 414
Qy      139  PheLeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAsp 158
Db      415  CTCCTGACGGCCACCCCGGTGAGTCTGTCAGCCGCTCCGCGAGGAGGAGGTGGCGGC 474
Qy      159  AlaIleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyr 178
Db      475  CTCGTGCGGAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
Qy      179  GlyIleIleIleSerIleMetIleArgThrSerVal---GlyAsnCysLysGlnLysGlu 197
Db      535  AGGCTCGTCAACGAGCTGCTGATGAGTGTCTCCGTGCGCGCGCGCGCGCGCGCGGAC 594
Qy      198  ArgLeuLeuSerValAlaAspAlaValAsnGluAlaIleThrSerPheGlyThrAlaAsp 217
Db      595  GAGTTCCTCGGCGCGCTCCACGAGGCCAAGAACCCAGCTGCTGCTCACCCTCCCGGAC 654
Qy      218  AlaPheProThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArg 237
Db      655  CTCCTTCGCTCGTCCAGCTCGCGCGATGTGGCGCGCGCGCGCGCGCGCGCGCGCG 714
Qy      238  LeuHisGlnGluIleAspAspIleLeuGluIleLeuAsnGluHisLys-----254
Db      715  AGCCGGAAGAGGATAGAGCGGATCATAGACACATCGTCCGGGAGCAGCAGGAGGTACATG 774
Qy      255  -----AlaAsnLysProPheGluAlaAspAsnLeu 264
Db      775  GGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 834
Qy      265  MetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValProValThrAsnGlu 284
Db      835  CTCACGCTCTCTGCTAGGCTGCGAAGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCG 894
Qy      285  SerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSerLysAla 304
Db      895  ATCATCGTGTCTCTCTGTTGACATGTTTTCAGGGGGGTAGCGAGACATCAGCCACAGTG 954
Qy      305  ThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGlu 324
Db      955  ATGATCTCGATCATGGCAGAGCTTATTCGGTGCCCAAGAGTAAATGACAAAGTACAGCC 1014
Qy      325  GluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHisAspLeu 344
Db      1015  GAGTTCGCCCAAGCCCTCCAAGGGAGGTCCACAGTCACTGAGGATGATATTGTAGGCTA 1074
Qy      345  LysPhePheLysLeuValLysGluThrLeuArgLeuHis---ProProValValLeu 363
Db      1075  AACTATCTAAATAATGTTGATCAAGGAGACCTTACGATTTGCAATTGCTTCCCTCGGACCCCTCTCA 1134

```



```

Db      1006  GCGAAGGCACAAGCTGAGGTACGACAAAGCTGCTGTTGGGAAGACATAGAGATCACAGAGAAT 1065
Qy      340  ArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisPro 359
Db      1066  GATATTGTTGGCTGAGCTACCTCAAGATGGTGATCAAGAGGCCCTTCGAGCTCATTC 1125
Qy      360  ProValValLeuLeu---ProArgGluCysArgGluThrThrArgIleAspGlyTrpGlu 378
Db      1126  CCGGCTCCCTTGGCTGAACCTCGTAATGTCTGAGACAAACCAAGTCATAGGCTATGAT 1185
Qy      379  IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db      1186  ATACCAAAAGGCACATCTGTGTTGTGAATATGTGGCGGATATGTAGGAGCCCTAATTAC 1245
Qy      399  TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLys 418
Db      1246  TGGGAGGATCCGAAGAGTTTAAGCCAGAGAGTTTGAATAATTAATGTTAGTAAAG 1305
Qy      419  GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db      1306  GGTAAATAACTTTGAATTCCTTCATTTGGCTCTGGACGTAGAAATTTGTCGGGTATAAAC 1365
Qy      439  SerAlaIleThrAsnLeuGluThrValIleIleAsnLeuLeuTyThrHisPheAsnTrpGlu 458
Db      1366  CTTGGGCTAGCCAACTTGAGCTAGCATTTAGCTAGTCTTCTACCATTTTGTACGTGAAG 1425
Qy      459  LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeu 478
Db      1426  CTACCAATGAATGTTGCCCAAGACCTTGACATGCAGGAGACCCCGGAATAGTTGCA 1485
Qy      479  ArgLysLysIleAspLeuLysLeuIleProIle 489
Db      1486  GCTAAACTAACTACTCTAAATATATGTCCTGTT 1518

RESULT 9
CL962214
LOCUS      OsIPCC007387 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL962214
VERSION    CL962214.1 GI:52379162
KEYWORDS   GSS.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM    Oryza sativa (indica cultivar-group)
REFERENCE   1 (bases 1 to 1536)
AUTHORS     Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE       An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
JOURNAL     Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES
            Location/Qualifiers
            1..1536
            /organism="Oryza sativa (indica cultivar-group)"
            /mol_type="genomic DNA"
            /db_xref="taxon:39946"
            /clone_lib="Oryza sativa Express Library"
            /note="Oryza sativa exon trapped genomic sequences "
ORIGIN
Alignment Scores:

```

```

Pred. No.:      2.13e-93      Length:      1536
Score:          910.00      Matches:      195
Percent Similarity: 57.0%      Conservative: 91
Best Local Similarity: 38.8%      Mismatches: 184
Query Match:    35.4%      Indels:      32
Db:             13      Gaps:      8

US-10-759-813-2 (1-500) x CL962214 (1-1536)

Qy      12  LeuIleSerPheLeuLeuValLeuIleLeuValVal-----23
Db      31  TTGATATCTTCTTCTCTCTCTCAAGCTCATCGCCGCTATGCTCGCTCGCGCAGG 90
Qy      24  -----ValMetArgLeuTrpLysGlnAsnProProGlyProTrpLysPhePro 41
Db      91  GAGAGTGGCTGCGGCTG-----CGGCCGGGCGCTCGAGCTGCGC 132
Qy      42  IleIleGlyAsnLeuProHisLeuLeuLeuThr-----SerAspLeuGlyHisGluArg 59
Db      133  CTGATCGGACGCTGACCACTCTCTCTGTACGGTACGGCGACCTCGCCGACCGGGCC 192
Qy      60  PheArgAlaLeuAlaGlnIleTyArgProValMetSerLeuGlnIleGlyGlnValSer 79
Db      193  ATGCGCGAGCTGCTCCGTACGCTACGGCCCTCATGCTCTCTCGGCTCGGCGCGTGC 252
Qy      80  AlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAla 99
Db      253  ACGCTGGTGTCTCTCTCGCGAGCGCGCGAGGTTCATGAGGCGCGACGACCGCCG 312
Qy      100  PheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyArgLysAspVal 119
Db      313  TTCGCGGCGCCACCTCAGCGCCACCATCGACATCTCTACGTGCGGCGGCGAGCATC 372
Qy      120  LeuPheAlaSerTyArgHisTrpArgGlnMetLysLysIleTrpIleLeuGluPhe 139
Db      373  ATCTTCGGGCGCTACACCGAGCGGTGGCGGAGCTCCGCAAGGTGTGCGGCTCGAGCTC 432
Qy      140  LeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAla 159
Db      433  TTCAACACCGCGCGTCTCTCTCTCGCGCCGCTCAGGAGGAGGAGGTGGG-----486
Qy      160  IleThrPheLeuArgSerLysAlaGlySerProVal-----Aen 172
Db      487  ---CGACTCTCTCCGCTCGCTCTCGCCGCTCCGCGAGGCGGCGCGCTGCTTCAAC 543
Qy      173  IleThrLysIleIleTyArgIleIleIleSerIleMetIleArgThrSerValGly---191
Db      544  CTCAGCGAGAGGATCTGCGCGCATGACCAACAGCTCCGTGTGTGCGCGCGCTCGCGCG 603
Qy      192  AsnCysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsnGluAlaThr 211
Db      604  CGGTGCGACCAACCGCGGAGGTTCCTCGACGAGCTCGACAAGCGCGGTGCGGCTCACCGGC 663
Qy      212  SerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHisTyIleIleGlyAla 231
Db      664  GGGATCAACTTGGCGGACCTGTACCCGCTCTCGGCTGTGGCGGCTCAGCGCGGCG 723
Qy      232  GluSerLysProArgArgLeuHisGlnGluIleAspAspIleLeuGluGluIleLeuAsn 251
Db      724  ACGCGGACATGGCGCGTTCGACAGGAACATCTACCGCATCGCGGAGCATCATCCGC 783
Qy      252  GluHisLysAlaAsnLysProPheGluAlaAsp-----AsnLeuMetAspValLeuLeu 269
Db      784  GACCGCGACGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 843
Qy      270  AsnLeuGlnLysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSer 289
Db      844  CGGCTGCGAGAGGCGCGGCGCTCAAGTTCGCACCTCACCCAGACATCATCAGCACCGTC 903
Qy      290  ValLeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMet 309
Db      904  ATTTTCGACATCTTCTCGCGCGGAGCGAGCGTCTGTCGACGACGCTAGACTGGACCATG 963

```



```
Db      814 GAGCCATGGCTGGCAGCCGCATCTTGATGATTTGGCCATCGAGGAGACCTGTGGAC 873
Qy      267 ValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValProValThrAsnGlnSerIle 286
Db      874 ACGCTCCTCAGGATCCAGAAGGAAGACACCTCGAGTGCCTCTCACACCGGCAACATC 933
Qy      287 LysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGlu 306
Db      934 AAAGCCGTTCTCTCGACATCTTTGGTCGAGAGTGATACATCTTCCACATCGTCCGAC 993
Qy      307 TrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluVal 326
Db      994 TGGGCTTGTGAGAGCTAATGAGAACCCAGAGGCGATGCACAAAGCACAAACCGAGCTT 1053
Qy      327 ArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhe 346
Db      1054 CGGAGTACCTCTCAAGGGAAGCAGATGGTAAGTGAAGACGACTTTGCAAGCCGTGACGTAC 1113
Qy      347 PheLysLeuValValLysGluThrLeuArgLeuHisProProVal---ValLeuIlePro 365
Db      1114 CTAAAGCTTGTTCATCAAGGAGACCCGAGGCTTCATCCCATGTGCCATTTGCTCTCCCG 1173
Qy      366 ArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIle 385
Db      1174 AGGAGTGTCTCAACATGCAAGGTATGGGTAGGTAGTGCCTCAAGGACACACCGTG 1233
Qy      386 ValValAsnAlaTrpAlaIleGlyArgAspProAsnThrThrTpsSerGluProGlyLysPhe 405
Db      1234 TTCTGTTAAGCTATGGCGATTAAACAGAGACCCCTAGGCACCTGGGATGAGCCTGAGGTTC 1293
Qy      406 AnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuVal 425
Db      1294 AAAGCCGAACGGTTTTCACATGCGAAGATCGACTTTAAAGGTGCAAAATTTCCAGTATATA 1353
Qy      426 ProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGlu 445
Db      1354 CGTTTGTGCGGTAGGAGGATTTGCCCTGGCATGACATTCGGACATGCTACCTGGAG 1413
Qy      446 TyrValIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrPro 465
Db      1414 CTCATGCTTGCAATGCTCTGTACCACTTCGACTGGGAGCTCCCTAAAGGAGTTGCACCG 1473
Qy      466 GlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLysLysIleAspLeuLys 485
Db      1474 AAGCAACTAGACATGACAGAGAGATGGGCATCACCGTGGTGGTAGGAAGAATGCTCTTTAT 1533
Qy      486 LeuIleProIleProTyrGlnValSerLeu 495
Db      1534 CTGCATCCTATA---GTCCGTGTGTCCTTG 1560
```

```
RESULT 11
LOCUS      CL965809
DEFINITION OaIFCC012786 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION  CL965809
VERSION     GSS.
KEYWORDS   CL965809.1 GI:52386299
SOURCE     Oryza sativa (indica cultivar-group)
ORGANISM   Oryza sativa (indica cultivar-group)
```

```
REFERENCE 1 (bases 1 to 1530)
AUTHORS   Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
```

```
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzaceae; Oryza.
```

```
FEATURES             Location/Qualifiers
     source            1..1530
     /organism="Oryza sativa (indica cultivar-group)"
     /mol_type="genomic DNA"
     /db_xref="taxon:39946"
     /clone_lib="Oryza sativa Express Library"
     /note="Oryza sativa exon trapped genomic sequences"
```

ORIGIN

```
Alignment Scores:
Pred. No.:      4,71e-93      Length:      1530
Score:          907.00      Matches:      190
Percent Similarity: 59.2%      Conservative: 109
Best Local Similarity: 37.6%      Mismatches: 182
Query Match:      35.2%      Indels:      24
DB:              13      Gaps:      8

US-10-759-813-2 (1-500) x CL965809 (1-1530)

Qy      12 LeuLeSerPheLeuLeuValLeuLeuValValMetArgLeuTrpLysLysGln 31
Db      13 CTCATGAGTACTTGTCTCTCTCGCTTCTCGTACCGCTCTTACCTCGCGCG 72
Qy      32 AsnPro-----ProProGlyProTrpLysPheProIle 42
Db      73 TCGTGTCTCGGAGGTTCGGGAGGTGCGGCTGCGCGGCGCGGCTGCGGCTGCGGTG 132
Qy      43 IleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAla 62
Db      133 ATCGGCCACCTGACACCACTCGCTCGCGCGCGCGCCACGACCGCGCATCGCGCAC 192
Qy      63 LeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValVal 82
Db      193 ATGCGCGCGCCACGCGCCGCTCATGCTGTCTCGGCTTCTCGAGCTCCCGGTGTGGTG 252
Qy      83 IleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGln 102
Db      253 GCCTCTCTCGCGGACGCCCGCGGAGATCATGAGGACGACGACGCTCGCGTTCGCGTCG 312
Qy      103 ArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAla 122
Db      313 CGGCGGATCGGCGCGATGCTCGGCTCGTTCAGAGGCGCGGAGGCGGTTCATCTTCGCG 372
Qy      123 SerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAla 142
Db      373 CCTACGCGCAGCGGTGGAGGACGCTCCGCAAGATCTGCACCGCTCGAGCTCTCAGCCAC 432
Qy      143 LysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPhe 162
Db      433 CGCGCGTCCACTCTCTCGCGCCCGCTCGCGCGGACGAGCTCGGCGCGCTCTCTCGCGCG 492
Qy      163 LeuArgSerLysAla-----GlySerProValAsnIleThrLysIleIleTyr 178
Db      493 GTGGCGGATCAGGCGCGCTTGTCTGTCGCTCTCGCGGTGAACCTCACCGGATGATCTCC 552
Qy      179 GlyIleIleIleSerIleMetIleArgThrSerValGlyAsn---CysLysGlnLysGlu 197
Db      553 GCCTTCGTCGGGACTCCACGCTGCGGCCCATCATTCGCGAGCGCGGAGGACCGCCGAC 612
Qy      198 ArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAsp 217
Db      613 ACGTCTCTCGCGTAGTGGAGACGGGCTCATGATCATGCGCGGAGTGGCTTCGCGGAC 672
Qy      218 AlaPheProThrTrpLysLeuHisTyrIleIleGlyAlaGluSerLysProArgArg 237
Db      673 CTGTTCCCTCTGTCGCGCTCGCCATGCTCTCTCAGCCGCGGTGCCCGCAAGATCGAGCGC 732
```

```
Db      814 GAGCCATGGCTGGCAGCCGCATCTTGATGATTTGGCCATCGAGGAGACCTGTGGAC 873
```

```
Qy      267 ValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValProValThrAsnGlnSerIle 286
```

```
Db      874 ACGCTCCTCAGGATCCAGAAGGAAGACACCTCGAGTGCCTCTCACACCGGCAACATC 933
```

```
Qy      287 LysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGlu 306
```

```
Db      934 AAAGCCGTTCTCTCGACATCTTTGGTCGAGAGTGATACATCTTCCACATCGTCCGAC 993
```

```
Qy      307 TrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluVal 326
```

```
Db      994 TGGGCTTGTGAGAGCTAATGAGAACCCAGAGGCGATGCACAAAGCACAAACCGAGCTT 1053
```

```
Qy      327 ArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhe 346
```

```
Db      1054 CGGAGTACCTCTCAAGGGAAGCAGATGGTAAGTGAAGACGACTTTGCAAGCCGTGACGTAC 1113
```

```
Qy      347 PheLysLeuValValLysGluThrLeuArgLeuHisProProVal---ValLeuIlePro 365
```

```
Db      1114 CTAAAGCTTGTTCATCAAGGAGACCCGAGGCTTCATCCCATGTGCCATTTGCTCTCCCG 1173
```

```
Qy      366 ArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIle 385
```

```
Db      1174 AGGAGTGTCTCAACATGCAAGGTATGGGTAGGTAGTGCCTCAAGGACACACCGTG 1233
```

```
Qy      386 ValValAsnAlaTrpAlaIleGlyArgAspProAsnThrThrTpsSerGluProGlyLysPhe 405
```

```
Db      1234 TTCTGTTAAGCTATGGCGATTAAACAGAGACCCCTAGGCACCTGGGATGAGCCTGAGGTTC 1293
```

```
Qy      406 AnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuVal 425
```

```
Db      1294 AAAGCCGAACGGTTTTCACATGCGAAGATCGACTTTAAAGGTGCAAAATTTCCAGTATATA 1353
```

```
Qy      426 ProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGlu 445
```

```
Db      1354 CGTTTGTGCGGTAGGAGGATTTGCCCTGGCATGACATTCGGACATGCTACCTGGAG 1413
```

```
Qy      446 TyrValIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrPro 465
```

```
Db      1414 CTCATGCTTGCAATGCTCTGTACCACTTCGACTGGGAGCTCCCTAAAGGAGTTGCACCG 1473
```

```
Qy      466 GlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLysLysIleAspLeuLys 485
```

```
Db      1474 AAGCAACTAGACATGACAGAGAGATGGGCATCACCGTGGTGGTAGGAAGAATGCTCTTTAT 1533
```

```
Qy      486 LeuIleProIleProTyrGlnValSerLeu 495
```

```
Db      1534 CTGCATCCTATA---GTCCGTGTGTCCTTG 1560
```



```
QY 238 LeuHieGlnLeuLeuAspPheLeuGluGluLeuAsnGluHisLysAlaAsnLys 257
Db 733 CGCGCGCGCGGATGATGGCTTCATCGACCATCATCCAGGACACCCAGGAGACGAGA 792
QY 258 ---ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnGluHisLysAsnGlyAsn 276
Db 793 GCGCGCGCGGACGAGGACCTTCGACGCTCTTGGAGGTCGAGAAAGACATGGAC 852
QY 277 ValProValProValThraAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAla 296
Db 853 TCCAGCATCCCTCACCACCATGAACATCAATCATCTTGATCAGATGTTCCGGCGG 912
QY 297 GlySerGluThrThrSerLysAlaThrGluThrValMetAlaGluLeuMetLysAsnPro 316
Db 913 GCGAGCGAGACGTCAGCGACGACGCTGAGTGGCGGATGCGGAGCTGATGCGTAACCG 972
QY 317 ThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGly----- 334
Db 973 GCGGTGATGCGCGCGGCGAGGAGGTGCGCGGAGAGCTCGCGCTCGCGCGGACGAC 1032
QY 335 LysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThr 354
Db 1033 AGGTTAACCGAGACACCTCCGAGCCTCCACTCTCGGCTGTGTCATCAGGAGAGCG 1092
QY 355 LeuArgLeuHisProProVal---ValLeuIleProArgGluCysArgGluThrThrArg 373
Db 1093 CTCGGCTCCACCGCGCAGCGCGCTCTCTCGCGCGGAGTGCGCGCGCGCTGCAAG 1152
QY 374 IleAspGlyThrGluIleHisProAsnThrArgIleValValAsnAlaThrPalaIleGly 393
Db 1153 GTGTTCCGCTACGATGCGCGCGGACGAGGTGCGCGGAGAGCTCGTGAACGCTGGCGATCGCG 1212
QY 394 ArgAspProAsnThrTrpSerGluProGly---LysPheAsnProGluArgPheLysAsp 412
Db 1213 CGGAGCGCGCGCTGGCGCGCGCGGAGAGGTCTCGCGGAGAGTTTCGAGCGG 1272
QY 413 CysAlaIleAspTrpLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArg 432
Db 1273 TCGGAGAGGACTTCAGGGGGCGGACTTCGAGCTCATCCGTTTCGCGCGCGGAGGAG 1332
QY 433 IleCysProGlyIleThrSerAlaIleThrAsnLeuLysIleValIleIleAsnLeu 452
Db 1333 ATCTGCCCGGATGCGCTTCGGGTGCGCGACGTCGAGCTCGCGCTCGCGCGCTGCTG 1392
QY 453 TyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGlu 472
Db 1393 TTCCTTCGACTGGAGGTGCGGAGGAGTGGCGCGCGGAGTGGAGATGAGATGAGCGAG 1452
QY 473 AlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeu-----Ile 487
Db 1453 GCGCGCGGATCAGGTGCGCGCGCTTCGGACCTACTGTTGCTGCGCGTTCGCGCGTGC 1512
QY 488 ProIleProThrGln 492
Db 1513 CCGGTGCCAGCGAG 1527
```

RESULT 12

CL965071

LOCUS

CL965071.1 2259 bp DNA linear GSS 21-SEP-2004
Osr1CC011624 Oryza sativa Expressed Sequence Tag (EST) library (indica
cultivar-group) genomic, genomic survey sequence.

DEFINITION

CL965071

ACCESSION

CL965071

VERSION

CL965071.1

KEYWORDS

GSS.

SOURCE

Oryza sativa (indica cultivar-group)

ORGANISM

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BPP
clade; Ehrhartoideae; Oryzoideae; Oryza.

REFERENCE

1 (bases 1 to 2259)

AUTHORS

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.

TITLE

JOURNAL

COMMENT

An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..2259

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Expressed Library"

/note="Oryza sativa exon trapped genomic sequences"

FEATURES

source

Alignment Scores:

Pred. No.: 1.45e-92 Length: 2259

Score: 905.00 Matches: 191

Percent Similarity: 57.9% Conservative: 102

Best Local Similarity: 37.7% Mismatches: 195

Query Match: 35.2% Indels: 18

DB: 13 Gaps: 8

US-10-759-813-2 (1-500) x CL965071 (1-2259)

ORIGIN

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuLeuSerPheLeuLeuValLeuIle 20

Db 730 CTGAACAGAGAAATGAGGAGCAGTATCTTCTTGGCCACTCTACTGATCTATCTCTC 789

QY 21 LeuValValMetArgLeuTrpLysLys---GlnAsnProProGlyProGlyProTrpLys 39

Db 790 GCCTTCGTCAAGCTCAGGCCACGCAACATGCGAGAGACCCGCTCCGGGCGCATCGCAG 849

QY 40 PheProIleIleGlyAsnLeuProHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 59

Db 850 CTGCGCGGTATCGGAGCGCTGCACAC-----CTGCGCGCGCGCTTCTCTACCGCGCG 903

QY 60 PheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSer 79

Db 904 CTGCGGTGACCTTGCCTCGCGCGCGCGCGCGAGCTCATGCTCGCGTCCGCGAGCTCC 963

QY 80 AlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAla 99

Db 964 GTGCTGCTGGCTGCTGCGCGCGCGCGCGAGGAGTGTAGGAGACCCACGACGCGCGCC 1023

QY 100 PheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspVal 119

Db 1024 TTCGCCACGCGCGCGCGAGCGCGCTCGCGCGCTCACCGCGAGCGCTCGCGCGCTC 1083

QY 120 LeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPhe 139

Db 1084 GCGTTCGCGCGCGAGGAGGAGCTGCGCTCCCTCCGAAAGCTCTGCTGCTCAGGAGCTC 1143

QY 140 LeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAla 159

Db 1144 CTCGCGCGCGCGCGCTCGCTGCTCGCGCGCGCGCGAGCGCGCGCGCGCGCTT 1203

QY 160 IleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGly 179

Db 1204 CTCGCGCTCGCTTTCACACGACGCGCGCGCGCGGAGCTGAGCTGCTGCTGCTGCTG 1263

QY 180 IleIleIleSerIleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArg 198

Db 1264 TACGTGACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1323

QY 199 LeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAla 218


```
Db 496 GTCCGCCGCGCGCGAGCGGCGGACCGGTGGAGCTCGCGGCGCGCGCTGTGCGCGCTC 555
Qy 181 llelleSerlleMetlleArgThrSerValGly---AsnCyLeYsGlnLysGluArgLeu 199
Db 556 GTGGCGGAGTCCACCGTGGCGCGCGTGTGGCGAGCGGTGGCGCGGCTCGACGTGTTC 615
Qy 200 LeuSerValAlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPhe 219
Db 616 CTCGCCAGCTCGACCGCGCCATCGAGTGTGTGGCTCAACGTGGCGGACCTGTGG 675
Qy 220 ProThrTrpLysLeuHieTyrllelleGlyAlaGluSerLysLysProArgArgLeuHis 239
Db 676 CGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 735
Qy 240 GlnGluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLysProPhe 259
Db 736 GACAGATGTTGGGTCTCGACGGCATCATCCAGCGCCACTGGAGAGAGTGGCGGC 795
Qy 260 GluAlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValProVal 279
Db 796 GCGGCGAAGACATTCAGCTGTCTACTAGGATACATAAGAGGCGGCGTGGAGTTC 855
Qy 280 ProValThrAsnGluSerlleLysAlaSerValLeuGlnMetPheThrAlaGlySerGlu 299
Db 856 CTTCTCGACATGAGCGCGTCAAAATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 915
Qy 300 ThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeu 319
Db 916 AGTGTGGCGACGACGCTGGGTGGCGTTCGCGGAGTAAATCCGGAACCGCGCGCGATG 975
Qy 320 ArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSer 339
Db 976 AAGAAGGCGACGCGGAGTTCGCGTGTGCTTTCGAGGCGCGCGCGTGTTCAGAGAGC 1035
Qy 340 ArgPhe-----HieAspLeuLysPheLysLeuValLysGluThrLeuArgLeu 357
Db 1036 GCGTGTGGCGTGGCGAGTCCCGTACCTTCGCTTCCCTGCTGTGTGTGTGTGTGTGTGT 1095
Qy 358 HisProProVal---ValleulleProArgGluCysArgGluThrThrArgileAspGly 376
Db 1096 CACCGCGCGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1155
Qy 377 TyrGluIleHisProAsnThrArgileValValAsnAlaTrpAlaIleGlyArgAspPro 396
Db 1156 TACGAGCTCCGCGCGCGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1215
Qy 397 AsnThrTrp-----SerGluProGlyLysPheAsnProGluArgPheLysAsp----- 412
Db 1216 CGGTACTGCGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1275
Qy 413 -----CysAlaIleAspTyrlLysGlyThrThrPheGluLeuValProPheGlyAlaGly 430
Db 1276 GCGCGCGCGCGGTGAGTTCAGGCGCGCGTTCGAGCTCTGCGGTTCGCGCGCGCGC 1335
Qy 431 LysArgIleCysProGlylleThrSerAlaIleThrAsnLeuGluTyrlVallelleAsn 450
Db 1336 CCGCGGATGTGCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1395
Qy 451 LeuLeuTyrlHisPheAsnTrpGluLeuAlaAspGlylleThrProGlnThrLeuAspMet 470
Db 1396 GTGTCTTCCATTCGCTGTGGAGGCTTCCGTTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1455
Qy 471 ThrGluAlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeulleProIle 489
Db 1456 ACAGAGGCGTTCGCGCATACGCGCGCGCAAGGCCAACCTCTGTGTGTGTGTGTGTGT 1512
```

```
RESULT 15
CL958628 1557 bp DNA linear GSS 21-SEP-2004
LOCUS OsIFCC001807 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL958628
VERSION CL958628.1 GI:52372005
```

```
KEYWORDS
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE 1 (bases 1 to 1557)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source 1..1557
Location/Qualifiers
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "
```

ORIGIN

```
Alignment Scores:
Pred. No.: 6.93e-92 Length: 1557
Score: 897.00 Matches: 189
Percent Similarity: 58.4% Conservative: 104
Best Local Similarity: 37.6% Mismatches: 183
Query Match: 34.8% Indels: 26
DB: 13 Gaps: 8
```

US-10-759-813-2 (1-500) x CL958628 (1-1557)

```
Qy 11 llellelleSerPheLeuValLeuValValMetArgLeuTrpLysLys 30
Db 22 CTACTTCTCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 81
Qy 31 GlnAsnPro-----ProProGlyProTrpTrpLysPheProIlelle 43
Db 82 CGGCATGCGCGCGCGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 141
Qy 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db 142 GGCACCTCTCCACCA-----GTCCGCGCGCGGTCTCCGCGCGCGCGCGCGCGCGCT 195
Qy 64 AlaGlnIleTyrlGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValle 83
Db 196 GCCCGCCCGCGCGCGCGCGCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 255
Qy 84 SerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 256 TGCACCGCGGAGCGCGCGGAGGTGACCAAGACCCACGACTCGCTTCGCGCGCGCG 315
Qy 104 ProIleValLeuAspAlaGlnIleValPheTyrlAsnArgLysAspValLeuPheAlaSer 123
Db 316 CCCATCACCCCGACCGGGAAGTCTCATGGCGGACAGCGTCCGCGGTGTGTGTGTGTGT 375
Qy 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db 376 TACGCGGAGGTTGAGAGACGCTCGCGAGGATCTGCACCTCGAGCTCTTAAGCGCGCG 435
Qy 144 LysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu 163
Db 436 CGCGTCCGCTCTCTCCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 495
```

Search completed: May 31, 2006, 08:36:43
Job time : 6934 secs

```
QY 164 ArgSerLysAlaGly-----SerProValAsnIle 173
Db 496 GCGCGCGCGCGCGGTGGCGGCTTGACAGACCCCGGGCGCGACGCGGTGAACCTG 555
QY 174 ThrLysIleIleTyrGlyIleIleIleSerIleMetIleArgThrSerValGlyAsn--- 192
Db 556 AGCGAGCGGATTTCCGGCGTACGTGGCGGACTCCGCGGTGCGGCCCTCATCTCGGAGCAGG 615
QY 193 CysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSer 212
Db 616 TTCAAGAACCGCGCGCTTCCTGCGGATCTGGAGCGGAGGATGAAGCTCTCCCGCGG 675
QY 213 PheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHisTyrIleIleGlyAlaGlu 232
Db 676 CAGTGTTCGCGGACCTTTCGCGTCTCGCGCGCGGATGCTCGTCAGCCGGATGCC 735
QY 233 SerLysProArgArgLeuHisGlnGluIleAspAspIleLeuGluIleLeuAsnGlu 252
Db 736 CGCCGGATGAAGCGGAGCGACAGGAGATGATGATTTCATCGACGACATCTTCCAGGAG 795
QY 253 HisLysAlaAsnLysProPheGluAla-----AspAsnLeuMetAspValLeuLeuAsn 270
Db 796 CATCAGAGAGCAGACCGCGCGCGCGCGGAGGAGACTGCTCATGCTCTCTTCGCGG 855
QY 271 LeuGlnLysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerVal 290
Db 856 ATCCAGAGCAAGACAAGACGAATCCCGCCCTCACCAACGACACATCAAGACCGGTGATC 915
QY 291 LeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaAlaThrGluTrpValMetAla 310
Db 916 ATCGACATGTTGTCGCGAGCAGCGAGACGCGCGCGGACGCTCGCTGCGAGTGGACCATGTCG 975
QY 311 GluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPhe 330
Db 976 GAGCTGATGAGGACCTAGGTGATGCGCAGGCGCAGGACGAGGTCCGACGGCCCTC 1035
QY 331 GlyLeuMetGlyLys-----ValAspGluSerArgPheHisAspLeuLysPhePheLys 348
Db 1036 GCCATTGCCGCGGAGCAGCGCGTCACCGAGGAGCGCTGCCGATCTCCCTTACCTGCAC 1095
QY 349 LeuValValLysGluThrLeuArgLeuHisProProVal---ValLeuIleProArgGlu 367
Db 1096 CTCCTCATCARGAGTCGTCGCGGTGCACCGCGGTGACCGATGCTGCTGCGGAGGAG 1155
QY 368 CysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIleValVal 387
Db 1156 TGCCGGAGCCGTGCGCGGTGATGGCTTCGAGTCCCGAGGGCGGTGATGTGCTCGTC 1215
QY 388 AsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnPro 407
Db 1216 AACGGGTGGCGATCGCAGGAGACCGCGGCACTGGGACTCGCGGAGGAGTTCCGCGCG 1275
QY 408 GluArgPheLys---AspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValPro 426
Db 1276 GAGAGTTCGAGGGCGCGCGCGCCCGCGACTTCAAGGGACCGGACTTCAGTACATACCG 1335
QY 427 PheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyr 446
Db 1336 TTCCGGCGCGCGCGGATGCCCCGGGATGGCGTTCGGGCTGGCCAAACATGGAGCTC 1395
QY 447 ValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGln 466
Db 1396 GCGCTCGCGCGCTCTCTGTACCATTCGACTGGGAGCTGCCCGCGGGATGCTTCCCGGC 1455
QY 467 ThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLysLysIleAspLeuLysLeu 486
Db 1456 GAGCTGGACATACGAGCGGCTTGGGGCTCACACGCGCGGTGCTCCGACCTCTCTGTT 1515
QY 487 IlePro 488
Db 1516 GTCCCT 1521
```

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:03:06 ; Search time 913 Seconds
(without alignments)
5727.477 Million cell updates/sec

Title: US-10-759-813-2

Perfect score: 2574

Sequence: 1 MEQKNLSPSILISFLVLI.....KIDKLPIPYQVSLGSLNIS 500

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/abs/ABSWEB spool/US10759813/runat_31052006_060302_2579/app_query.fasta_1
-DB=N_Geneseq -QFM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05h
-USER=US10759813 @CN 1 1 761 @runat 31052006_060302_2579 -NCPU=6 -ICPU=3
-NO WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV -TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.8:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*
15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2574	100.0	1733	6	ABK11137
2	1164	45.2	1538	12	ADN10568 Nicotiana
3	1164	45.2	1538	14	ADZ65256 Tobacco f

4	1164	45.2	1538	14	ADZ65803	Tobacco c
5	1164	45.2	1538	14	AEE06775	Tobacco p
6	1164	45.2	1538	15	AEE06775	Tobacco c
7	1149	44.6	1673	12	ADN10646	Nicotiana
8	1149	44.6	1673	12	ADN10626	Nicotiana
9	1149	44.6	1673	14	ADZ65334	Tobacco f
10	1149	44.6	1673	14	ADZ65314	Tobacco f
11	1149	44.6	1673	14	ADZ65861	Tobacco c
12	1149	44.6	1673	14	ADZ65881	Tobacco c
13	1149	44.6	1673	14	AEE06853	Tobacco p
14	1149	44.6	1673	14	AEE06833	Tobacco p
15	1149	44.6	1673	15	AEE06833	Tobacco c
16	1149	44.6	1673	15	AEE06833	Tobacco c
17	1148	44.6	1673	12	ADN10624	Nicotiana
18	1148	44.6	1673	14	ADZ65312	Tobacco f
19	1148	44.6	1673	14	ADZ65859	Tobacco c
20	1148	44.6	1673	14	AEE06831	Tobacco p
21	1148	44.6	1673	15	AEE06831	Tobacco c
22	1144	44.4	1657	12	ADN10610	Nicotiana
23	1144	44.4	1657	14	ADZ65298	Tobacco f
24	1144	44.4	1657	14	ADZ65845	Tobacco c
25	1144	44.4	1657	14	AEE06817	Tobacco p
26	1144	44.4	1657	15	AEE06817	Tobacco c
27	1142	44.4	1673	12	ADN10628	Nicotiana
28	1142	44.4	1673	14	ADZ65316	Tobacco f
29	1142	44.4	1673	14	ADZ65863	Tobacco c
30	1142	44.4	1673	14	AEE06835	Tobacco p
31	1142	44.4	1673	15	AEE06835	Tobacco c
32	1131	43.9	1610	12	ADN10676	Nicotiana
33	1131	43.9	1610	12	ADN10612	Nicotiana
34	1131	43.9	1610	14	ADZ65300	Tobacco f
35	1131	43.9	1610	14	ADZ65364	Tobacco f
36	1131	43.9	1610	14	ADZ65911	Tobacco c
37	1131	43.9	1610	14	ADZ65847	Tobacco c
38	1131	43.9	1610	14	AEE06819	Tobacco p
39	1131	43.9	1610	14	AEE06883	Tobacco p
40	1131	43.9	1610	15	AEE06883	Tobacco c
41	1131	43.9	1610	15	AEE06883	Tobacco c
42	1129.5	43.9	1576	12	ADN10564	Nicotiana
43	1129.5	43.9	1576	14	ADZ65252	Tobacco f
44	1129.5	43.9	1576	14	ADZ65799	Tobacco c
45	1129.5	43.9	1576	14	AEE06771	Tobacco p

ALIGNMENTS

RESULT 1
ABK11137
ID ABK11137 standard; cDNA; 1733 BP.
XX
AC ABK11137;
XX
DT 05-JUN-2002 (first entry)
XX
cDNA encoding Euphorbia lagascae cytochrome P450 enzyme.
XX
Cytochrome P450 enzyme; synthesis of delta 12-epoxy fatty acid; epoxide;
veronolic acid; modified fatty acid; oil; commercial crop; plasticiser;
crosslinking coating application; setting printing ink; transgenic;
plant; EST; clone eelc_pk002_14; gene; ss; expressed sequence tag.
XX
Euphorbia lagascae.
XX
Key Location/Qualifiers
CDS 14..1516
/*tag= a
/product= "Cytochrome P450 enzyme"
XX
WO200208269-A2.
XX
31-JAN-2002.
XX
19-JUL-2001; 2001WO-US022790.

```

XX PR 21-JUL-2000; 2000US-0219833P.
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI Cahoon EB;
XX WPI; 2002-257331/30.
XX DR P-PSDB; AAU77470.
XX PT Novel isolated polynucleotide encoding plant cytochrome P450 enzyme
XX PT associated with synthesis of Delta12-epoxy fatty acids, useful for
XX PT creating transgenic plants with higher or lower level expression of the
XX PT enzyme.
XX PS Claim 8; Page 49; 53pp; English.
XX CC The present invention relates to the isolation of the polynucleotide and
CC polypeptide sequences for a plant cytochrome P450 enzyme associated with
CC the synthesis of delta 12-epoxy fatty acids from Euphorbia lagascae. The
CC polynucleotide sequence of the invention can be used to produce epoxide
CC containing fatty acids such as vernolic acid. The sequences of the
CC invention can be used to manipulate modified fatty acids to produce oils
CC in commercial crops. They can also be used to produce plasticizers, for
CC crosslinking coating applications, and setting printing inks. The
CC polynucleotide can also be used for creating transgenic plants in which
CC the enzyme is present at higher or lower levels than normal, in cell
CC types or in developmental stages in which they are not normally found.
CC The present sequence encoding the E. lagascae cytochrome P450 enzyme of
CC the invention is isolated from the expressed sequence tag (EST) clone
CC sellic.pk002.14
XX SQ Sequence 1733 BP; 556 A; 282 C; 376 G; 519 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,566-238 Length: 1733
Score: 2574.00 Matches: 500
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-759-813-2 (1-500) x ABK1137 (1-1733)
QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
DB 14 ATGGAGCAGAGAAAATCTCTTTCCGAGCATTTTAAATAGTTTCTGCTGTTTAAATC 73
QY 21 LeuValValMetArgLeuThrLysLysGlnAsnProProProGlyProThrLysPhe 40
DB 74 TTAGTAGTAGTCATGAGGTTGTGGAGAAACAGAAATCCACCTCCAGGGCCATGGAAGTTT 133
QY 41 ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe 60
DB 134 CCTATCATAGGAAATCTCTCTCATTTATTTACTCATCTCTGATCTAGGCCATGAAGCTTTT 193
QY 61 ArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
DB 194 AGAGCCTTGGCTCAATTTATGACCTGTATGAGTCTTCAAAATGGCCAAAGTTTCAGCT 253
QY 81 ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe 100
DB 254 GTTGTCAATTTCTTTCAGCTGAAGCAGCAGAAAGAGTTATGAAACCTCAGGCTGATCCCTTC 313
QY 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu 120
DB 314 GCCCAACGCCCTTATCGTCTTGACCCACAGATGTGTTTATTAATCGGAAAGATGCTCTG 373
QY 121 PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
DB 374 TTTGCTTCATATGGAGATCCTGGAGCAGATGAAGAAATTTGGATACTTGAATTTCTG 433
QY 141 SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIle 160

```


ID	ADN10568	standard; cDNA; 1538 BP.
XX	AC	ADN10568;
XX	XX	15-JUL-2004 (first entry)
DT	XX	Nicotiana p450 enzyme encoding cDNA SEQ ID NO:153.
XX	XX	plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
KW	KW	phenotype; gene; ss.
XX	OS	Nicotiana sp.
XX	XX	WO2004035745-A2.
FN	XX	29-APR-2004.
PD	XX	16-OCT-2003; 2003WO-US032722.
PF	XX	16-OCT-2002; 2002US-0418933P.
XX	PR	08-JUL-2003; 2003US-0485368P.
XX	PR	18-SEP-2003; 2003US-0503989P.
XX	XX	(USSM-) US SMOKELESS TOBACCO CO.
PA	XX	Claim 1; SEQ ID NO 153; 198pp; English.
PI	XX	Xu D;
XX	XX	WPI; 2004-348441/32.
DR	XX	P-PSDB; ADN10569.
XX	XX	New isolated nucleic acid molecules and encoded cytochrome P450 enzymes from Nicotiana plants, useful for altering plant phenotypes.
PT	XX	Claim 1; SEQ ID NO 153; 198pp; English.
PT	XX	The present sequence represents a plant p450 enzyme nucleic acid molecule isolated from Nicotiana. Also described: (i) an isolated p450 protein from Nicotiana; (2) a transgenic plant comprising the nucleic acid molecule described above; (3) a method of producing the transgenic plant, comprising: (i) operably linking the above nucleic acid molecule with a promoter functional in the plant to create a plant transformational vector; (ii) transforming the plant with the vector; (iii) selecting a plant cell transformed with the transformation vector; and (iv) regenerating a transformation plant from the transformed plant cell; and (4) a method of selecting a plant containing the above nucleic acid molecule, where the plant is analysed for the presence of the above nucleic acid sequences. The p450 sequences have plant growth regulant activity, and can be used in gene therapy. Compositions and methods from the present invention are useful for altering plant phenotypes.
XX	XX	Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;
XX	XX	Alignment Scores:
	Pred. No.:	4.5e-102 Length: 1538
	Score:	1164.00 Matches: 226
	Percent Similarity:	63.4% Conservative: 88
	Best Local Similarity:	45.7% Mismatches: 165
	Query Match:	45.2% Indels: 16
	DB:	12 Gaps: 5
US	10-759-813-2	(1-500) x ADN10568 (1-1538)
Qy	12	LeuIleSerPheLeuValLeuIleLeuValValMetArgLeuTrpIysGln 31
Db	48	TTGGTTCCATTTCCTATTCTTCCTTTCTTTTGTAAAGGAATGGAAGAACTCG 107
Qy	32	AsnPro-----ProProGlyProTrpLysPheProIleLeuGlyAenLeu 46
Db	108	AATAGCCCAAGAAAAAATGCCACCGTCCATGGAACTACCAATACTAGGAAGATG 167
Qy	47	ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db	168	CTTCATATGGTT-----GGTGGATACACACCATGTCCTTAGAGATTAGCCAAAAA 221

QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
 DB 1302 GTTGTAAATATTTGAATATCTTCATTTGGTGGAGGAGGATTTCCAGGGATT 1361
 QY 438 ThrSerAlaIleThrAsnLeuGluTrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
 DB 1362 TCGTTGGTTAGCTAAATGCTTATTTGGCATTGGCTCAATTACTTATCATCTTGATTGG 1421
 QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
 DB 1422 GAATCCCCACTGGGAATCAACACAGCAGCTTGGACTTGACTGAGTTGGTGGAGTAACT 1481
 QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
 DB 1482 GCCGCTAGAAAAGTGACCTTACTTGGTTGGGACTCCTCTATCAA 1526
 RESULT 3
 ID ADZ65256 standard; cDNA; 1538 BP.
 XX
 AC ADZ65256;
 XX
 XX 14-JUL-2005 (first entry)
 XX Tobacco full length cDNA for cytochrome p450 clone D90A-BB3.
 DE
 DE Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
 KW normocotine; transgenic plant; gene.
 XX
 XX Nicotiana tabacum.
 OS
 XX WO2005038033-A2.
 PN
 XX 28-APR-2005.
 PD
 XX 15-OCT-2004; 2004WO-US034065.
 PF
 XX 16-OCT-2003; 2003US-00686947.
 PR 29-APR-2004; 2004US-05662359.
 PR 03-SEP-2004; 2004US-00934944.
 XX
 XX (USSM-) US SMOKELESS TOBACCO CO.
 PA
 XX Xu D;
 PI
 XX WPI: 2005-315717/32.
 DR P-PSDB; ADZ65257.
 XX
 XX New nucleic acid molecule encoding cytochrome P450 enzymes in Nicotiana,
 PT useful in developing tobacco plants with altered phenotypes.
 XX
 XX Disclosure; SEQ ID NO 153; 226pp; English.
 PS
 XX The invention relates to an isolated nucleic acid molecule (I) from
 CC Nicotiana, where the nucleic acid molecule comprising any of the 59
 CC nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
 CC SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
 CC ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic
 CC acids of the invention encode cytochrome p450 enzymes whose expression is
 CC induced by ethylene and/or plant senescence. Also included are a
 CC transgenic plant comprising (I), a method of producing a transgenic
 CC plant, a method of selecting a plant containing a nucleic acid molecule
 CC (where the plant is analyzed for the presence of nucleic acid sequence of
 CC ADZ65402-
 XX
 XX Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 4.5e-102 Length: 1538
 Score: 1164.00 Matches: 226
 Percent Similarity: 63.4% Conservative: 88
 Best Local Similarity: 45.7% Mismatches: 165
 Query Match: 45.2% Indels: 16

DB: 14 Gaps: 5
 US-10-759-813-2 (1-500) x ADZ65256 (1-1538)
 QY 12 LeuIleSerPheLeuValIleLeuValValMetArgLeuTriPlysLysGln 31
 DB 48 TTGGTTCCATTTTCCTATTTCTCTTTCTTTTGGTAAATAGGAAGAACTCG 107
 QY 32 AsnPro-----ProProGlyProTriPlysPheProIleIleGlyAsnLeu 46
 DB 108 AATAGCCAAAGGAAAAATTCGCCACGAGTCCATGGAAACTACCAATACTAGGAAGTATG 167
 QY 47 ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
 DB 168 CTTCATATGGTT-----GGTGGACTACCAACCATGTCCTTAGAGATTTTAGCCAAAAA 221
 QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
 DB 222 TATGACCGCTTATGACCTTCAATTAGGTGAGTTCTGAGTTGGTGTACTTCTCTCT 281
 QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
 DB 282 GATATGGCAAAAGAAAGTACTAAAAAATCATGACATCGCTTTTCGCGCTAGCGCTTAGCCTT 341
 QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
 DB 342 TTGGCCCCGGAGATTGCTGTACAATAGGTCTGATCTTGGCTTTTGGCCCTTATGGCGAT 401
 QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
 DB 402 TATTGGAGACAATGCCGTAAATATGCTCTTGGAGTGTCTCAGTCCCAAGAATGTCGG 461
 QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
 DB 462 ACATATAGCTCTATTAGCGCGATGAAGTCTCTCGTCTCTTAATTTTATCCGGTCATCT 521
 QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
 DB 522 TCTGGTGAGCCTGTTAATATTACGAAAGAGATCTTTTGTTCACAAGCTCCATGACATGT 581
 QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
 DB 582 AGATCAGCGTTGGCAAGTATTCAGGAGCAAGCAAAATTTATACAACATAATTAAGAA 641
 QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTriPlysLeuLeu 225
 DB 642 GTTATACTCTTAGCAGGAGGGTTGATGTGGCTGACATATTCCTTCAACAAGTCTCTT 701
 QY 226 HisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIle 245
 DB 702 CATGTGCTCAGTGGGAATGAAGGTAAAGATTATGAATGCACACCATGAAGTAGATGCTATT 761
 QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
 DB 762 GTTGAGATGTCATCAGCAGCACCAAGAAAATCTTGCAATTTGGGAAAACATAATGGAGCG 821
 QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
 DB 822 TTAGGAGGTGAAGATTAAATTTGATGTTCTTCTTAAAAACTTATGAATGATGGAGCGCTCAA 881
 QY 279 ValProValThrAsnGluSerIleLysValAserValLeuGlnMetPheThrAlaGlySer 298
 DB 882 TTTCTCTATCACAACGACCAACATCAAGCTATAATCTTTTGACATGTTTGTCTGCTGGAACA 941
 QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
 DB 942 GAGACTTCATCGTCAACAATTTGTGGGCTATGGTGAATAATGGTGAATAATCCAACGTGA 1001
 QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
 DB 1002 TTTGCGAAAGCTCAAGCAGAAAGTAAAGATGCAATTATTAGAGAAAAAGAAACTTTTGTGATA 1061
 QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358

Db 1062 AATGATGTGGAGAGCTAAATCTATCTAAAGTTAGTCTAAAGAAACTCTAAGACTTCAT 1121
Qy 359 ProProVal---ValleuileProArgGluCysAtgGluThrThrArgileAspGlyTyr 377
Db 1122 CCACCGTTCCACTTTGCTCCAGAGAAATGTAGGGAAGAGACAAATATAACGGCTAC 1181
Qy 378 GluileHisProAsnThrArgileValValAsnAlaTrpAlalleGlyArgAspProAsn 397
Db 1182 ACTATTCTCTGTAAGACCAAGTCATGGTTAATGTTTGGCATTTGGGAAGAGATCCAAA 1241
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlalleAspTyr 417
Db 1242 TATTGGGATGATCGCAAACTTTTAAGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTT 1301
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgileCysProGlylle 437
Db 1302 GTTGGTAATAATTTTGAATATCTTCCATTTGGTGGGAGGAGGATTTGCCAGGAT 1361
Qy 438 ThrSerAlaileThrAsnLeuGluTyrValilleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1362 TCGTTTGGTTAGCTAATGCTTATTGTCATTTGGCTCAATTACTTTATCTATTGATGG 1421
Qy 458 GluLeuAlaAspGlylleThrProGlnThrLeuAspMetThrGluAlalleGlyAla 477
Db 1422 GAATCCCTCCACTGGAATCAACCAAGCGACTTGGACTTGAAGTGGTGGAGTAAC 1481
Qy 478 LeuArgLysLysIleAsnLeuLysLeuIleProIleProTyrGln 492
Db 1482 GCCGTAGAAAAAGTGACCTTTACTTGGTTGGGACTCTCTTATCAA 1526

RESULT 4

ID AD265803 standard; cDNA; 1538 BP.

XX AD265803;

XX 14-JUL-2005 (first entry)

XX Tobacco cytochrome P450 enzyme cDNA #77.

XX Enzyme engineering; cytochrome P450; gene; ss.

XX Nicotiana tabacum.

XX WO2005038018-A2.

XX 28-APR-2005.

XX 15-OCT-2004; 2004WO-US034218.

XX 16-OCT-2003; 2003US-00686947.

XX 29-APR-2004; 2004US-0568235P.

XX 17-SEP-2004; 2004US-00943507.

XX (USSM-) US SMOKELESS TOBACCO CO.

XX Xu D;

XX WPI; 2005-315709/32.

XX P-PSDB; AD265804.

XX New isolated nucleic acid molecule from Nicotiana, useful for altering plant phenotypes, thus producing a transgenic plant having reduced levels of nicotine.

XX Example 3; SEQ ID NO 153; 203bp; English.

XX The invention relates to an isolated nucleic acid molecule from

XX Nicotiana, encoding a protein. The invention also relates to a transgenic plant comprising the nucleic acid molecule, a method of producing a transgenic plant comprising operably linking the nucleic acid molecule with a promoter functional in the plant to create a plant

CC transformational vector, transforming the plant with the plant transformational vector, selecting a plant cell transformation with the transformational vector and regenerating a transformation plant from the transformed plant cell, a method of selecting a plant containing a nucleic acid molecule, a method of increasing or decreasing nicotine levels in a plant by operably linking the nucleic acid molecule with a promoter functional in the plant, a tobacco product having reduced amounts of nicotine in the plant, the tobacco product comprising tobacco from the plant, a tobacco leaf having reduced amounts of nicotine levels and a method of isolating a gene from a plant using the isolated nucleic acid. In producing a transgenic plant, the plant has reduced levels of nicotine. The tobacco product is selected from cigarettes, cigars, pipe tobacco, snuff, chewing tobacco, products blended with tobacco product and their mixtures. The nucleic acid molecule is useful for altering plant phenotypes, thus producing a transgenic plant having reduced levels of nicotine. This sequence represents cDNA encoding a tobacco cytochrome P450 enzyme of the invention.

XX Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,5e-102 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
DB: 14 Gaps: 5

US-10-759-813-2 (1-500) x AD265803 (1-1538)

Qy 12 LeuileSerPheLeuLeuValleuileLeuValValMetArgLeuTrpLysGln 31
Db 48 TTGGTTTCCATTTCTTATCTTTCTTTCTTTTGTAAAGAAATGGAAGACTCG 107
Qy 32 AsnPro-----ProProGlyProTrpLysPheProIleleGlyAsnLeu 46
Db 108 AATAGCCAAAGGAAAAAATTTGCCACGAGTCCAGAACTACCACTAGTAGAATG 167
Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 168 CTTTATATGGTT-----GGTGAGCTACACACCATGCTCTTAGAGATTAGCCAAAAA 221
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 222 TATGACCGCTTATGCACCTTCAATTAGTGAAGTTTCTGCAGTTGTGTACTTCTCT 281
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 282 GATATGGCAAAAGAGAGTACTAAAAAATCATGACATCGCTTTTCGCGCTAGGCTT 341
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 342 TTGGCCCGGAGATTGCTGTACAAATAGTCTGTATCTTTCGCTTTGCCCTTATGCGGAT 401
Qy 127 HistTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGAGACAAATCGTAAATAATATGTGCTTTGGAAGTCTCAGTCCCAAGAAATGTCG 461
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 462 ACATATAGCTCTATTAGGCGCGATGAAGTTCTTCGTCTCTTAATTTATCCGGTCTAT 521
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 522 TCTGTGAGCTCTTAATATATTACGGAAGAGATCTTTTGTTCACAGCTCCATGACATGT 581
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 582 AGATCAGCGTTTGGGCAAGTATTCAAGGAGCAAGACAAATTTATACAACATAATTAAGAA 641
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 642 GTTATACTCTTAGCAGGAGGGTTTGTATGTGGTGCACATATATCCCTTCATCAAGTCTCT 701

Qy	226	HistyrIlelleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAsePile	245
Db	702	CATGTGCTAGTGAATAAGGGTAAGATTATGAATCACACCATATAGGTAGATGCTATT	761
Qy	246	LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro	258
Db	762	GTTTGAGNATGTCATCAGCAGCACAGAAAAATCTTCCAATTGGGAAAACACTATGAGCG	821
Qy	259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
Db	822	TTAGGAGGTGAAGATTTAAATTGATGTCTCTCAAAACTTATGAATGATGAGGCCTCAA	881
Qy	279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	882	TTTCCTATCACCAACGACAACATCAAGCTATAATCTTTGCATGTTTGTCTGTGNACA	941
Qy	299	GluThrThrSerTyAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu	318
Db	942	GAGACTTATCTCGTCAACAATTTGTGTGGCTATGTTGGAATGGTGAATAAATCCAACTGT	1001
Qy	319	LeuArgLysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValaspGlu	338
Db	1002	TTTCGAAAGCTCAAGCAGAAGTAAGATGCAATTTTAGAGAAAAGAAACTTTTGATGAA	1061
Qy	339	SerArgPheHisAspLeuLysPhePheLysLeuValVallysGluThrLeuArgLeuHis	358
Db	1062	AATGNTGTGGAGGAGCTAAACTATCTAAAGTTAGTCAATTAAGAAACTCTAAGACTTCAT	1121
Qy	359	ProProVal---ValLeuIleProArgGluCyArgGluThrThrArgIleAspGlyTyr	377
Db	1122	CCACCGGTTCCACTTTTGCTCCCAGAGAATGTAGGGAAGAGACAAATATAAACGGCTAC	1181
Qy	378	GluIleHisProAsnThrArgIleValValAsnAlaTrpAlalleGlyArgAspProAsn	397
Db	1182	ACTATTCTGTAAGACCMAAGTCATGTTAAATGTTTGGGCATTTGGGAGAGATCCAAA	1241
Qy	398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlalleAspTyr	417
Db	1242	TATTGGGATGATGCAGAAACTTTTAAGCCAGAGAGATTTGAGCAGTCTCTTAAGGATTTT	1301
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Db	1302	GTTTGTAAATTTTGAATATCTTCCATTTTGGTGGTGAAGGAGATTGTCACGGGATT	1361
Qy	438	ThrSerAlalleThrAsnLeuGluTyrValleIleAsnLeuLeuTyrHisPheAsnTrp	457
Db	1362	TCGTTTGGTTTTAGCTAATGCTTATTTGCCCATTTGGCTTGAATGCTTGGTTGGTGA	1421
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlalleGlyGlyAla	477
Db	1422	GAATCCCCACTGGAAATCAACACCAAGCGACTTGGACTTGAATGCTTGGTTGGATGA	1481
Qy	478	LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln	492
Db	1482	GCCGCTAGAAAAAGTGACCTTTACTTGTTCGCACTCTTATCAA	1526

RESULT 5

RESULT 3
AEE06775

AEE06775
ID AEE06775 standard; cDNA: 1538 BP.

1D AEE06775
XX
AC AEE06775;

AC
XX
DT

DE Tobacco p450 cDNA SEQ ID NO: 162.

XX
KW

KW transgenic plant; senescence.

XX
OS Nicotiana tabacum.

XX
PN WO2005111217-A2.

[illegible]

24-NOV-2005.

27-APR-2005; 2005WO-US014803.

29-APR-2004; 2004US-0566235P.

03-SEP-2004; 2004US-00934944.
03-SEP-2004; 2004US-0607357P.

17-SEP-2004; 2004US-00943507.
15-OCT-2004; 2004WO-US034065.

15-OCT-2004; 2004WO-US034218.
25-JAN-2005; 2005US-0646764P.

24-MAR-2005; 2005US-0665097P.
24-MAR-2005; 2005US-0665451P.

(USSM-) US SMOKELESS TOBACCO

Xu D, Nielsen MT;

WPI; 2005-786788/80.

P-PSDB; AEE06776.

Producing a tobacco plant having decreased expression of a nicotine demethylase gene comprises crossing a first tobacco plant with a second tobacco plant and germinating the collected seed of an F1 progeny plant.

Claim 32: SEQ ID NO 162: 641pp; English.

The invention relates to a breeding method for producing a tobacco plant with reduced expression of a nicotine demethylase gene comprising crossing a first tobacco plant with variant nicotine demethylase gene expression with a second tobacco plant with at least one phenotypic trait to produce an F₁ progeny plant, the seed of which is collected and germinated to produce a tobacco plant having reduced expression of a nicotine demethylase gene. Also included are breeding a nicotine demethylase demethylase gene. Also included are breeding a tobacco demethylase gene.

deficiency trait into a tobacco plant, producing a tobacco seed, developing a tobacco plant in a tobacco breeding program, a tissue culture of regenerable tobacco cells obtained from the tobacco plant of the invention, producing a tobacco product, a breeding method for producing a tobacco plant having a modified attribute, a method of breeding an attribute into a tobacco plant, a tobacco plant or its components produced by the method of breeding a nicotine demethylase deficiency trait into a tobacco plant, producing tobacco seed, producing a tobacco plant having a modified attribute or developing a tobacco plant in a tobacco breeding program, an isolated genetic marker comprising a nucleic acid sequence that is substantially identical to a nucleic acid sequence given in the specification (the nucleic acids comprise isolated cytochrome p450 cDNAs), an expression vector comprising the nucleic acid sequence, a plant or plant component comprising the isolated nucleic acid sequence, a plant produced from a germinated seed of the plant, reducing the expression or enzymatic activity of a constitutive, or an ethylene induced or senescence induced tobacco polypeptide in a plant cell, and increasing the expression or enzymatic activity of a constitutive, or an ethylene or senescence induced tobacco polypeptide in a plant cell. The phenotypic trait comprises disease resistance, high yield, high grade index, curability, curing quality, mechanical harshness, holding ability, leaf quality, height, maturation, stalk size, or leaf number per plant. The breeding method for producing a tobacco plant having decreased expression of a nicotine demethylase gene is useful developing desirable (non-genetically engineered) germplasm. The plant is useful in producing (smokeless) tobacco products. The tobacco product is a moist or dry snuff, a chewing tobacco, a cigarette product, a cigar product, a cigarillo, a pipe tobacco, or bidis. The p450 cDNAs were isolated using degenerate PCR primers designed against cytochrome p450 motifs. The present sequence is a cytochrome p450 cDNA of the invention.

Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Alignment Scores:

Alignment Scores:	
Pred. No.:	4.5e-102
Score:	1164.00
Percent Similarity:	63.4%
Best Local Similarity:	45.7%
Conservative:	88
Mismatches:	165
Length:	1538
Matches:	226

Query Match:	45.2%	Indels:	16
DB:	14	Gaps:	5
US-10-759-813-2 (1-500) x ABE06775 (1-1538)			
Qy	12	LeuIleSerPheLeuLeuValValLeuLeuValValMetArgLeuTrpLysGln	31
Db	48	TTGGTTTCATTTTCCTATTTCTATCTTTCTTTTTTGTAAAGAAATGGAAGAACTCG	107
Qy	32	AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu	46
Db	108	AATAGCCAAAGAAAAAATTGCCACGATCCATCGAAACTACCAATACTAGAAAGTAGT	167
Qy	47	ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	168	CTTCATATGGTT-----GGTGGATACACACCATGCTCTTAGAGATTAGCCAAAAA	221
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla	86
Db	222	TATGGACCGCTATGCACCTTCATAGGTGAAGTTTCTGCAGTTGTGTACTTCTCTCT	281
Qy	87	GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal	106
Db	282	GATATGGCAAAAGAGTAGTACTAAAACTCATGACATCGCTTTTCGCGCTAGGCCCT	341
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126
Db	342	TTGCCCGCGAGATTGTCGTTACAAATAGGTCGATCTTGCGTTTGGCCCTATGCGCAT	401
Qy	127	HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln	146
Db	402	TATGGAGACAAATCGGTAAATATGTCGTTGGAAAGTGCTCAGTGCCCAAGATGTTCCG	461
Qy	147	SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys	166
Db	462	ACATATAGCTCTATTAGCGCGCATGAAGTCTTCGTCTCTCTTAATTTTTCGGTCACT	521
Qy	167	AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle	186
Db	522	TCTGGTGAGCCTGTTAATTACGGAAAGGATCTTTTGTTCACAAAGTCCATGACATGT	581
Qy	187	ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla	205
Db	582	AGATCAGCGTTTGGCAAGTATTCAAGGAGCAAGACAAATTTATACACTTAATTAAGAA	641
Qy	206	ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu	225
Db	642	GTTATACTCTTAGCAGGAGGGTTTGATGGCTGACATATTCCTTCATACAAAGTCTCT	701
Qy	226	HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle	245
Db	702	CATGTGCTCAGTGAATGAAGGTAAGATTATGAATGCACACCATAAGGTAGATGCTATT	761
Qy	246	LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro	258
Db	762	GTTGAGAAATGTCATCAACGAGCACAGAAAAATCTTGCAATTGGGAAACTAATGGAGCG	821
Qy	259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
Db	822	TTAGGAGGTGAAGATTAAATTGATGTTCTCTTAAAACTTATCAATGATGGAGGCCCTCA	881
Qy	279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	882	TTTCTCTATCACACGACCAATCAAGCTAATATCTTTTGACATGTTTGCTGCTGGAACA	941
Qy	299	GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu	318
Db	942	GAGACTTCATCGTCAACAAATTTGTGGGCTATGGTGGAAATGGTGAAAAATCCAACCTA	1001
Qy	319	LeuArgLysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGlu	338
Db	1002	TTTCGAAAGCTCAAGCAGAAGTAGAGATGCATTTAGAGAAAAAGAACTTTTGATCAA	1061

Qy	339	SerArgPheHisAspLeuLeuPhePheLeuValVallysGluThrLeuArgLeuHis	358
Db	1062	AATGATGTGGAGGAGCTAAACTATCTAAAGTTAGTCAATTAAGAAACTCTAAAGACTTCAT	1121
Qy	359	ProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgIleAspGlyTyr	377
Db	1122	CCACCGGTTCACATTTTGCTCCCAAGAGAATGTAGGGAAGAGACAATAATAAACGGCTAC	1181
Qy	378	GluIleHisProAsnThrArgIleValValAsnAlaTirpAlaIleGlyAArgAspProAsn	397
Db	1182	ACTATTCTGTAAAGACCAAAAGTCATGTTAATGTTTGGGCATTTGGGAAGAGATCCAAA	1241
Qy	398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
Db	1242	TATTGGGATGATGCAGAAACTTTTAAGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTT	1301
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Db	1302	GTTGGTAATAATTTTGAATATCTTCCATTTGGTGGTGAAGGAGGATTTGTCCAGGGATT	1361
Qy	438	ThrSerAlaIleThrAsnLeuGluTyr=ValIleIleAsnLeuLeuTyrHisPheAsnTirp	457
Db	1362	TCGTTTGGTTAGCTAATGCTTATTGCCATTGGCTCAATTACTTTATCACTTTGATTGG	1421
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
Db	1422	GAACCTCCCACTGGAATCAAACCAAGCGGACTTGGACTTGACTGAGTTGGTGGAGTAACT	1481
Qy	478	LeuArgIlyLysIleAspLeuLeuLysLeuIleProIleProTyrGln	492
Db	1482	GCCCGTAGAAAAAGTGACCTTTACTTTGGTTGCGACTCCTTATCAA	1526
RESULT 6			
AEF97168	ID AEF97168 standard; cDNA; 1538 BP.		
XX	AC	AEF97168;	
XX	DT	20-APR-2006 (first entry)	
XX	DE	Tobacco cytochrome P450 cDNA clone D90A-BB3, SEQ ID NO:162.	
XX	KW	Plant breeding; crop improvement; secondary metabolite; genetic marker;	
XX	KW	plant; cytochrome P450; CYP; gene; ss.	
XX	OS	Nicotiana tabacum.	
XX	XX	US2006041949-A1.	
XX	PD	23-FEB-2006.	
XX	XX	27-APR-2005; 2005US-00116881.	
XX	XX	13-NOV-2001; 2001US-0337684P.	
XX	XX	11-JAN-2002; 2002US-0347444P.	
XX	XX	12-MAR-2002; 2002US-0363684P.	
XX	XX	16-OCT-2002; 2002US-0418933P.	
XX	XX	13-NOV-2002; 2002US-00293252.	
XX	XX	10-JAN-2003; 2003US-00340861.	
XX	XX	12-MAR-2003; 2003US-00387346.	
XX	XX	08-JUL-2003; 2003US-0485368P.	
XX	XX	18-SEP-2003; 2003US-0503989P.	
XX	XX	16-OCT-2003; 2003US-00686947.	
XX	XX	29-APR-2004; 2004US-0566235P.	
XX	XX	03-SEP-2004; 2004US-00934944.	
XX	XX	03-SEP-2004; 2004US-0607357P.	
XX	XX	17-SEP-2004; 2004US-00943507.	
XX	XX	15-OCT-2004; 2004WO-US034065.	
XX	XX	15-OCT-2004; 2004WO-US034218.	
XX	XX	25-JAN-2005; 2005US-0646764P.	
XX	XX	24-MAR-2005; 2005US-0665097P.	
XX	XX	24-MAR-2005; 2005US-0665451P.	
XX	XX	19-APR-2005; 2005US-00110062.	

```

XX PA (USSM-) US SMOKELESS TOBACCO CO.
XX PI Xu D, Nielsen MT;
XX WP1; 2006-183895/19.
XX P-P5DB; AEF97169.
DR New breeding method, useful for producing a tobacco plant having
PT decreased expression of a nicotine demethylase gene comprises crossing
PT germinating F1 progeny seed to produce the tobacco plant.
XX Claim 63; SEQ ID NO 162; 51lpp; English.
XX PS
CC The invention relates to a breeding method for producing a tobacco plant
CC having decreased expression of a nicotine demethylase gene. The method
CC involves crossing a first tobacco plant having variant nicotine
CC demethylase gene expression with a second tobacco plant containing at
CC least one phenotypic trait (e.g., disease resistance, high yield etc.) to
CC produce an F1 progeny plant; collecting the seed of the F1 progeny; and
CC germinating the seed to produce a tobacco plant having decreased nicotine
CC demethylase expression. The invention also relates to a tobacco plant or
CC its components produced using the method of the invention; a tissue
CC culture of regenerable tobacco cells obtained from such plants; a tobacco
CC product produced from such plants; and a method of breeding a tobacco
CC plant with a modified attribute comprising variant expression of a
CC cytochrome P450 polynucleotide (including the nicotine demethylase gene).
CC The invention further relates to isolated constitutive, ethylene-induced
CC or senescence-induced genetic markers comprising tobacco cytochrome P450
CC nucleic acid molecules, including nicotine demethylase sequences; methods
CC for reducing or increasing the expression or activity of polypeptides
CC encoded by these nucleic acid molecules; expression vectors, plants or
CC plant components comprising one of these polynucleotides; and tobacco
CC products produced from such plants. The methods of the invention are
CC useful for the breeding (especially marker assisted breeding) of tobacco
CC plants with decreased nicotine demethylase expression or with altered
CC cytochrome P450 expression or activity. Altered expression of such
CC enzymes can result in a change in the composition of secondary
CC metabolites such as alkaloids (e.g., nicotine), phenylpropanoids,
CC terpenoids, lipids, cyanogenic glycosides and glucosinolates, with
CC effects on the flavor or aroma of plant products. They may also affect
CC herbicide tolerance, resistance to disease or insects, quality factors
CC related to undesirable constituents, structural traits, fiber content,
CC leaf yield, ripening, leaf curing or storage properties. Tobacco plants
CC of the invention in which the expression of cytochrome P450 genes is
CC altered may have desirable traits such as altered levels of nicotine
CC or N'-nitrososnicotine. Such plants can be used in the production of
CC tobacco products such as moist or dry snuff, chewing tobacco, cigarettes,
CC cigars, cigarillos, pipe tobacco, bidis or smokeless tobacco products.
CC The present sequence represents a specifically claimed tobacco cytochrome
CC P450 polynucleotide useful in marker assisted breeding methods. Note: The
CC sequence data for this patent is also available in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=20060041949.
XX SQ Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservatives: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
DB: 15 Gaps: 5

US-10-759-813-2 (1-500) x AEF97168 (1-1538)
QY 12 LeuIleSerPheLeuValLeuValValMetArgLeuTrpLysGln 31
Db 48 TTGGTTTCATTTCTCTATCTTTCTCTTTCTTTCTTTTAAAGAAATGGAAGAACTCG 107
QY 32 AanPro-----ProGlyProTrpLysPheProIleIleGlyAenLeu 46
Db 11

108 AATAGCCAAAGGAAAAAATTCACCAGGTCCTCCAGGAACTACCAATACTAGGAATATG 167
47 ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
168 CTTTCATATGGTT-----GGTGGACTACCAACACCATGTCCTTAGAGATTTAGCCAAAAA 221
67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
222 TATGACCGCTTATGACCTTTCAATTAGGTGAAGTTCTTCGAGTTGTGCTTCTCTCT 281
87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
282 GATATGGCAAAAGAAAGTACTAAAACTCATGACATCGCTTTCGGCTCTAGGCTAGCCTT 341
107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
342 TTGGCCCGGAGATTGCTGTGTACATAGGTCTGATCTTGGCTTTTGCCCTATGCGCAT 401
127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
402 TATTGGAGACAAATGCGTAAATATGTCTTGGAAAGTCTCAGTGCACCAAGAAATGTCGG 461
147 SerSerArgLeuIleAtgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
462 ACATATAGCTTATTAGGCGCATGAAGTCTTCGTCCTCTTAATTTATCCGGTCACT 521
167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
522 TCTGGTGAAGCTGTTAATATTACGAAAGGATCTTTTGTTCACAAAGCTCCATGATGT 581
187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
582 AGATCAGCGTTGGCGCAAGTATTCAAGAGGACAAACAAATTTATACAACTAATTAAGAA 641
206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
642 GTTATCTCTTAGCAGGAGGTTTGATGTGCTGCATATTTCCCTTCATCAAGTCTCTT 701
226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle 245
702 CATGTGCTCAGTGAAGTGAAGGTAAGATTATGAATGCACACCATTAAGTAGATGCTATT 761
246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
762 GTTGAGATGTCATCAAGCAGCAGCAAGAAAAATCTTGCAATTTGGAAAACTAATGAGCG 821
259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlnValPro 278
822 TTAGAGGTGAAGATTAAATTGATGTTCTCTTAAACCTTATGATGATGAGGCGCTTCAA 881
279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
882 TTTCTCTATCACAACGACACATCAAGCTATAATCTTTTGACATGTTTGTGCTGGAACA 941
299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
942 GAGACTTCATCGTCAACAAATTGTGGCTATGTGGAAATGGTGAANAATCCAATCTGTA 1001
319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
1002 TTTGCGAAGCTCAAGCAGAGAGTAAGAGATGCAATTTAGAGAAAAAGAAACTTTTTCATGAA 1061
339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
1062 AATGATGTGGAGAGCTAAACTATCTAAAGTTAGTTCATTAAAGAAACTCTAAGACTTCAT 1121
359 ProProVal----ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
1122 CCACGGTTCACCTTTTCTCCCAAGAAATGTAGGAAAGACAAATATAAACCCTGCTAC 1181
378 GluIleHisProAsnThrArgIleValValAsnAlaIleGlyArgAspProAsn 397
1182 ACTATTCTCTAAAGACCAAAAGTCATGTTTAATGTTTGGGCATTTGGGAAGAGATCAAAA 1241

```

```
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1242 TATTGGGATGATCGACAGAACTTTTAAGCCAGAGAGATTGACAGTGTCTTAAGGATTTT 1301
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgLysCysProGlyIle 437
Db 1302 GTTGGTAAATATTTGAATATCTTCCATTTGGTGGTGAAGGAGGATTTGTCAGGAT 1361
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1362 TCTTTGGTTTACCTAATGCTTATTGTCCTATGCTCAATTAATCTTATCATTGATGG 1421
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1422 GAATCCCCACTGGAATCAACCAAGCAGCTTGGACTTGACTGAGTGGTGGAGTAAC 1481
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1482 GCGGCTAGAAAAAGTGACCTTTACTTGGTTGGCACTCCTTATCAA 1526
```

RESULT 7

ADN10646

ID ADN10646 standard; cDNA; 1673 BP.

XX ADN10646;

XX AC

XX 15-JUL-2004 (first entry)

XX Nicotiana p450 enzyme encoding cDNA SEQ ID NO:231.

XX plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;

XX phenotype; gene; ss.

XX Nicotiana sp.

XX WO2004035745-A2.

XX 29-APR-2004.

XX 16-OCT-2003; 2003WO-US032722.

XX 16-OCT-2002; 2002US-0418933P.

XX 08-JUL-2003; 2003US-0485368P.

XX 18-SEP-2003; 2003US-0503989P.

XX (USSM-) US SMOKELESS TOBACCO CO.

XX Xu D;

XX WPI; 2004-348441/32.

XX P-PSDB; ADN10647.

XX New isolated nucleic acid molecules and encoded cytochrome P450 enzymes from Nicotiana plants, useful for altering plant phenotypes.

XX Claim 1; SEQ ID NO 231; 198pp; English.

XX The present sequence represents a plant p450 enzyme nucleic acid molecule isolated from Nicotiana. Also described: (i) an isolated p450 protein from Nicotiana; (2) a transgenic plant comprising the nucleic acid molecule described above; (3) a method of producing the transgenic plant, comprising: (i) operably linking the above nucleic acid molecule with a promoter functional in the plant to create a plant transformational vector; (ii) transforming the plant with the vector; (iii) selecting a plant cell transformed with the transformation vector; and (iv) regenerating a transformation plant from the transformed plant cell; and (4) a method of selecting a plant containing the above nucleic acid molecule, where the plant is analysed for the presence of the above nucleic acid sequences. The p450 sequences have plant growth regulant activity, and can be used in gene therapy. Compositions and methods from the present invention are useful for altering plant phenotypes.

SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Alignment Scores:

```
Score: 1.41e-100 Length: 1673
Pred. No.: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: 12 Gaps: 5
```

US-10-759-813-2 (1-500) x ADN10646 (1-1673)

```
QY 12 LeuIleSerPheLeuLeuValLeuValMetArgLeuTrpLysGln 31
Db 54 TTGGTTTCCATTTCTATTTCTTCCTCTTTTGTAAAGATATGAAGAAGACTCC 113
QY 32 AsnPro-----ProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 AATAGCCAAAGCAAAAGTTGCCACAGGTCCATGGAACTACCAATACTAGGAAGTATG 173
QY 47 ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuGlnIle 66
Db 174 CTTTCATGGTT-----GGTGGACTACCAACCATGTCTTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGGACCACTTATGCACCTTCATTTAGGTGAAGTTTCGCGGTGTGGTTACTTCTCT 287
QY 87 GluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATCGGCAAAAGAGATATTAATAACTCATGACATCGCTTTTGGCGTCTAGGCGCTAGCCTT 347
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCGGAGATTGTCTGTACAAATAGGTCTGTAGTCTAGCCCTTGGCCCTATGGCGAC 407
QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTCGAGACAAATCGGTAAATATATGTCTTTGGAAGTCTCAGTGCACCAAGATGTCGG 467
QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGGCGGAATGAAGTCTCTCGTCTCATTAATTTTATCCGGTCATCT 527
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
Db 528 TCTGTGAACCTATTAAATGTACGAAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCAGAGCAGCAAGCAAAATTTATACACTAATTAAGAA 647
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGGTTTCATGTGGTGCACATATTCCTTCTACTGAAGTTCTT 707
QY 226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnLeuIleAspIle 245
Db 708 CATGTGCTCAGTGGAAATGAAGGTAAAGATTATGAATGCACACCAAGGTAGTACCAT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCAGCAAGAAATCTTGCATTTGGGAAACTAATGGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTAATTAATGATGTTCTTCTTAAGACTTATGAATCATGAGGCGCTCAA 887
QY 279 ValProValThrAsnGluSerIleLysValAspValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTATCACCAGCAACATCAAGCCATAATTTTTTGACATGTTTGTGCGCGGAC 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
```

```
Db 948 GAGACTTCATCGTCACCAATTGTGGGCTATGGTAGAAATGGTGAAAAATCCAGCCGTA 1007
Qy 319 LeuArgLysAlaGlnGluGlnValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGGCAAGCTCAGCAGAGTAGAGAGCATTTAGAGGAAAAGAACTTTCGATGAA 1067
Qy 339 SerArgPheHisAspLysPhePheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGCTAACTACCTACCTAAGTTAGTAAATAAAGAACTCTAAGACTTCAT 1127
Qy 359 ProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgLysAspGlyTyr 377
Db 1128 CCACGGGTTCCACTTTGCTCCCAAGAGAAATGTAGGGAAGACAAATATATAACGGCTAC 1187
Qy 378 GluIleHisProAsnThrArgIleValValAenAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCCTGTAAAGACCAAGCTCATGTTAATGTTTGGCTTTGGGAAGAGATCCAAA 1247
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAAACTTTTATGTCAGAGAGATTTGAGCAGTGCTCTAAGGATTTT 1307
Qy 418 LysGlyThrThrPheGluValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGGTAATAATTTGAATATCTCCATTTGGTGGGGAAGGAGATTGTCCTGGGATT 1367
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTTAATGCTTATTGGCATTGGCTCAATTACTATATCACTTCGATTGG 1427
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1428 AAATCCCTCGTGGAAATCGAACCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAAC 1487
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGCACTCCTCTTATCAA 1532
RESULT 8
ADN10626
ID ADN10626 standard; cDNA; 1673 BP.
XX
AC ADN10626;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nicotiana p450 enzyme encoding cDNA SEQ ID NO:211.
XX
KW plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
KW phenotype; gene; ss.
XX
OS Nicotiana sp.
XX
PN WO2004035745-A2.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-US032722.
XX
PR 16-OCT-2002; 2002US-0418933P.
PR 08-JUL-2003; 2003US-0485368P.
PR 18-SEP-2003; 2003US-0503989P.
XX
PA (US5M-) US SMOKELESS TOBACCO CO.
XX
PI Xu D;
XX
XX WPI; 2004-348441/32.
DR P-PSDB; ADN10627.
XX
PT New isolated nucleic acid molecules and encoded cytochrome p450 enzymes
from Nicotiana plants, useful for altering plant phenotypes.
```

```
XX Claim 1; SEQ ID NO 211; 198pp; English.
XX
CC The present sequence represents a plant p450 enzyme nucleic acid molecule
isolated from Nicotiana. Also described: (i) an isolated p450 protein
from Nicotiana; (2) a transgenic plant comprising the nucleic acid
molecule described above; (3) a method of producing the transgenic plant,
comprising: (i) operably linking the above nucleic acid molecule with a
promoter functional in the plant to create a plant transformational
vector; (ii) transforming the plant with the vector; (iii) selecting a
plant cell transformed with the transformation vector; and (iv)
regenerating a transformation plant from the transformed plant cell; and
(4) a method of selecting a plant containing the above nucleic acid
molecule, where the plant is analysed for the presence of the above
nucleic acid sequences. The p450 sequences have plant growth regulant
activity, and can be used in gene therapy. Compositions and methods from
the present invention are useful for altering plant phenotypes.
XX
SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.41e-100 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.8% Indels: 16
DB: 12 Gaps: 5
```

US-10-759-813-2 (1-500) x ADN10626 (1-1673)

```
Qy 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysGln 31
Db 54 TTGTTTTCCATTTTCCTATTTCTATCTTCCTTTTGGTAAAGATATGAAGAATCC 113
Qy 32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 AATAGCCAAAGCAAAAGTTGCCACCAGTCCATGGAACTACCAATACTAGGAAGTATG 173
Qy 47 ProHisLeuLeuLeuThrSerAspLysGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCATATGGTT-----GGTGGACTACCAACCACCTGCTTAGAGATTTAGCCAAAAA 227
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGACCACCTTATGCACCTTCAATAGGTGAAGTTTCTCGGTTGCGTTACTTCTCTCT 287
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAAAGAAAGTATTAAAAAATCATGACATCGCTTTTGGCTTAGGCCTAGCCTT 347
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCCGGAGATTGCTGTGTACAATAGGTCTGTATCTAGCCTTTTGGCCCTATGCGCAC 407
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATCGTAAATATGCTCTTGGAGTGTCTCAGTCCCAAGAAATGTTCCG 467
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTCGTCTCATTAATTTATCCGGTCATCT 527
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGGTGAACCTATTAAATGTTACGAAAGAGATCTTTTGTTCACAAGTCCCATGATGT 587
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCAGAGCAAGCAAAATTTATACAACTAATTAAGAA 647
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGGTTTATGTGCTGACATATTCCCTTCACTGAAGTTTCTT 707
```


Db 228 TATGGACCACTTATGACCTTCAATTAGGTGAAGTTCTCGCGTTGGTTACTTCTCCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAGAAGATATAAAGTCTATGACATCGCTTTTCGCTAGGCTAGCCTT 347
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCGGAGATTGCTGTGTACATAGTCTGATCTAGCCTTTTGCCCTATGCGAC 407
QY 127 HisTyrArgGlnMetLysLysIleTyrPheLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATGCGTAAATATATGCTCTTGGAAAGTCTCAGTCCCAAGAATGTCG 467
QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTAGTCTTATTAGCGGAATGAAGTTCTTCGTCTCATTAATTTATCCGGTCTCT 527
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
Db 528 TCTGGTGAACCTATTATGTTACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGCTTGGGCAAGTGTCAAAGAGCAAGCAAAATTTATACAACTAATTAAGAA 647
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTTGTAGCAGGAGGGTTGATGGGTGCATATTCCTTCACTGAAGTTTCTT 707
QY 226 HisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspIle 245
Db 708 CATGTGCTCAGTGGATGAAGGTAAAGATTATGAATGCACACCAATAAGTAGATGCCAT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGAATGTCATCAATGAGCACAAAGAAATCTTGAATTTGGGAAACTAATGAGCG 827
QY 259 PheGluAlaAspAsnMetAspValLeuLeuAsnLeuGlnLysAsnValPro 278
Db 828 TTAGAGGTGAAGATTAAATGATGTTCTTCAAGACTTATGAATGATGAGCGCTCAA 887
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCCTATCACCACCAACATCAAGCCATAATTTTGACATGTTTCTGCTCCGGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTCATCTCAACAATTTGTGGGCTATGTTAGTAATGTTGAAATAATCCAGCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCCGGAAGCTCAAGCAAGTAAGAGCAAGCATTTAGAGGAAAGAAACTTTTCATGAA 1067
QY 339 SerArgPheHisAspLysPhePheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1068 ATATGATGGAGGAGCTAACTACCTAAGTTAGTAATAAAGAAACTTAGACTTCTAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTGCTCCCAAGAGAATGTAGGGAAGACAAATATATAACGGCTAC 1187
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCCTGTAAGACCAAGATCATGGTTAATGTTGGCTTTTGGGAAGAGATCCAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAGAACTTTTATGCCAGAGAGATTGACAGCTGCTCTAAGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437

Db 1308 GTTGGTAATAAATTTTGAATATCTTCATTTGGTGGCGAAGGAGGATTTGTCTCTGGATT 1367
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTyr 457
Db 1368 TCGTTGGCTTAGCTAATGCTTATTTCGCATTGGCTCAATTACTATATCATCTTCGATTGG 1427
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AACTCCCTGCTGGATCGAATCGAACCAAGCAGCTTGGACTTGACTGAGTTGGTGGAGTAACT 1487
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTACCTTTACTTGGTTGGAGCTCCTTATCAA 1532
RESULT 11
ADZ65861
ID ADZ65861 standard; cDNA; 1673 BP.
XX
AC ADZ65861;
XX
DT 14-JUL-2005 (first entry)
XX
DE Tobacco cytochrome P450 enzyme cDNA #106.
XX
KW Enzyme engineering; cytochrome P450; gene; ss.
XX
OS Nicotiana tabacum.
XX
PN WO2005038018-A2.
XX
PD 28-APR-2005.
XX
PF 15-OCT-2004; 2004WO-US034218.
XX
PR 16-OCT-2003; 2003US-00686947.
PR 29-APR-2004; 2004US-0566235P.
PR 17-SEP-2004; 2004US-00943507.
XX
PA (USSM-) US SMOKELESS TOBACCO CO.
XX
XU Xu D;
PI
XX
DR WPI; 2005-315709/32.
XX
PT P-PSDB; ADZ65862.
XX
PT New isolated nucleic acid molecule from Nicotiana, useful for altering
XX plant phenotypes, thus producing a transgenic plant having reduced levels
XX of nornicotine.
PS
XX Disclosure; SEQ ID NO 211; 203pp; English.
CC
XX The invention relates to an isolated nucleic acid molecule from
CC Nicotiana, encoding a protein. The invention also relates to a transgenic
CC plant comprising the nucleic acid molecule, a method of producing a
CC transgenic plant comprising operably linking the nucleic acid molecule
CC with a promoter functional in the plant to create a plant
CC transformational vector, transforming the plant with the plant
CC transformational vector, selecting a plant cell transformed with the
CC transformational vector and regenerating a transformation plant from the
CC transformed plant cell, a method of selecting a plant containing a
CC nucleic acid molecule, a method of increasing or decreasing nornicotine
CC levels in a plant by operably linking the nucleic acid molecule with a
CC promoter functional in the plant, a tobacco product having reduced
CC amounts of nornicotine levels, the tobacco product comprising tobacco
CC from the plant, a tobacco leaf having reduced amounts of nornicotine
CC levels and a method of isolating a gene from a plant using the isolated
CC nucleic acid. In producing a transgenic plant, the plant has reduced
CC levels of nornicotine. The tobacco product is selected from cigarettes,
CC cigars, pipe tobacco, snuff, chewing tobacco, products blended with the
CC tobacco product and their mixtures. The nucleic acid molecule is useful
CC for altering plant phenotypes, thus producing a transgenic plant having
CC reduced levels of nornicotine. This sequence represents cDNA encoding a
CC tobacco cytochrome P450 enzyme of the invention.

XX SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,41e-100	Length:	1673
Score:	1149.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	44.6%	Indels:	16
DB:	14	Gaps:	5

US-10-759-813-2 (1-500) x ADZ65861 (1-1673)

Qy	12	LeuLeSerPheLeuLeuValValLeuLeLeuValValValMetArgLeuTrpLysLysGln	31
Db	54	TTGGTTTCCATTTCCTATTCCTCTCTTTTGTAAAGGATATGGAAGAACTCC	113
Qy	32	AsnPro-----ProProGlyProTrpLysPheProIleLeuGlyAsnLeu	46
Db	114	ANTAGCCAAAGCAAAAGTTGCCACAGGTCTCATGAAACTACCAATACTAGCAAGATG	173
Qy	47	ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	174	CTTCATATGGT-----GGTGGACTACACACCATCTCTTAGAGATTAGCCAAAAA	227
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla	86
Db	228	TATGGACCACCTTATGCACCTTCAATTAGGTGAAGTTTCTGCGGTTGTGTTACTTCTCT	287
Qy	87	GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal	106
Db	288	GATACGGCAAAAGAAGTATTAATAAACTCATGACATCGCTTTTGGCTTAGGCCTAGCCCT	347
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126
Db	348	TTGGCCCCGGAGATTGCTGTTTACAATAGTCTGATCTAGCCTTTTGGCCCTATGGCGAC	407
Qy	127	HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln	146
Db	408	TATTGGAGACAATCGTAAAAATATGTGTTCTGGAAAGTCTCAGTGCCAAAGATGTTGCG	467
Qy	147	SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys	166
Db	468	ACATTTAGCTCTATTAGCGGGAATGAAGTCTTCGTCTCATTAATTTATCCGTCATCT	527
Qy	167	AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle	186
Db	528	TCTGTGAACCTATTAAATGTTACGGAAGAGATCTTTTGTTCACAAGCTCCATGACATGT	587
Qy	187	ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla	205
Db	588	AGATCAGCGTTTGGCGCAAGTGTTCAAAGAGCAAGACAAAATTTATACAACTAATAAGAA	647
Qy	206	ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu	225
Db	648	GTGATACTTTAGCAGGAGGTTTGNATGGGCTGCATATTCCCTTCACCTGAAGTTCTT	707
Qy	226	HisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIle	245
Db	708	CATGTGCTCAGTGGAAATGAAGGTAAAGATTATGAATGCACACCATAAAGGTAGATGCCA	767
Qy	246	LeuGluGluIleLeuAsnGlnHis-----LysAlaAsnLysPro	258
Db	768	GTTGAGAATGTCATCAATGAGCACAAGAAAAATCTTGGCAATTGGGAAAAACTAATGAGCG	827
Qy	259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
Db	828	TTAGGAGGTGAAGATTAAATGATGTCTCTTCAAGCTTATGAATGATGGAGCCCTCAA	887
Qy	279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	888	TTTCTATCACCACGACAACTCAAGCAATAATTTTGGATGTTTGTGTCGCGGAC	947

528	Db	TCTGGTGAACCTATTAAATGTTACGGAAAGGATCTTTTGTGTCAAAAGCTCCATGACATGT	587
187	Qy	ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla	205
588	Db	AGATCAGCGTGTGGCGCAAGTGTTCAAAGACGACAGACAAATTTATACAACTTAATTAAGAA	647
206	Qy	ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu	225
648	Db	GTGATACTTTAGCAGGAGGGTTTGATGTGCGCTGCACATATCTCCCTCACATGAAGTTTC	707
226	Qy	HisTyrlleileGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle	245
708	Db	CATGTGCTCAGTGGAAATCAAGGGTAGATTATGAATGCCACACCATATAGGTAGATGCCATT	767
246	Qy	LeuGluGluLeuLeuAsnGluHis-----LysAlaAsnLysPro	258
768	Db	GTTGAGAAATGTCATCAATGATGACACAGAAAAATCTTGCATTTGGGAAAACTAATGGAGCG	827
259	Qy	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
828	Db	TTAGGAGGTGAAGATTTAATTGATGTCTTCTTAAGACTTATGAATGATGGAGGCCITCAA	887
279	Qy	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
888	Db	TTTCCTATACCAACGACCAACATCAAGGCCATAATTTTTCGATCATGTTTGTCTGCCGGGACA	947

XX DT 26-JAN-2006 (first entry)


```
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
Db 228 TATGGACCACCTTATGCACCTTCATAGTAGTGAAGTTTCTGCGGTGTGGTTACTTCTCT 287
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAGAAGTAGTATTAATAAACTCATGACATCGCTTTTGGCGCTAGGCCTT 347
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAsnSerTyrGlyAsp 126
Db 348 TTGGCCCGGAGATTGCTGTACAAATAGGCTGATCTAGCTTTTGGCCCTTAGCGCAC 407
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATCGTAAATAATGTGCTTGGAAAGTCTCAGTGCACGAAGAATGTCGG 467
Qy 147 SerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGGAATGAAGTTCTTCGTCTCATTAATTTTATCCGGTCATCT 527
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGTGTGAACCTATTATGTGTACGAAGAGATCTTTTGTTCACAAAGCTCCATGACATGT 587
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCAAAGAGCAAGCAAAATTTATACAACTAATTAAGAA 647
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGAGGGTTTGATGTGGCTGACATATTCCTTCTCAAGTTTCTT 707
Qy 226 HisTyrIleIleGlyAlaGlnSerLysProArgLeuHisGlnGluIleAspAspIle 245
Db 708 CATGTGCTCAGTGAATGAGGTGAATGATTATGAATGCACACCATAGGTAGTATGCCATT 767
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAAGAAAATCTTGCATTTGGGAAAACATAATGGAGG 827
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGAGAGTGAAGATTAAATTCATGTCTTCTTAAGACTTATGAATGATGGAGCGCTTCAA 887
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTCTATCACCACCAACACATCAAGCCATAATTTTGGACATGTTTGTCTCCCGGACA 947
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTCATCGTCAACATTTGTGTGGCTATGTTAGAAATGTTGTAATAATCCAGCCGTA 1007
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGCGAAGCTCAAGCAAGATGACAGAGCAATTTAGAGGAAAAAGAACTTTTCATGAA 1067
Qy 339 SerArgPheHisAspLeuLysPheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGAGCTAACTACCTAAAGTTAGTATTAATAAAGAACTTAAGACTTCAT 1127
Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTTGTCTCCCAAGAGAAATGTAGGGAAGACAAATATAAACGGCTAC 1187
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCTCTGTAAAGACCAAACTGATGTGTATGTTTGGGCTTTGGGAAGAGATCAAAA 1247
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGATGACGACAGAACTTTTATGCCAGAGAGATTTGACGAGTGTCTTAAGGATTTT 1307
```

```
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGCTAATAATTTTGAATATATCTTCCATTGTCGCGAAGAGGATTGTCTCTGGGATT 1367
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTAATGCTATTGTCATTTGGCTCAATTACTATATCACTTCGATGG 1427
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AAATCCCTCTGCTGGAATCGAACCAAGCGACTTGGACTTGAAGTTGGTTGGAGTAACT 1487
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGTGTCGACTCCTTATCAA 1532
RESULT 15
AEF97226
ID AEF97226 standard; cDNA; 1673 BP.
XX AEF97226;
XX 20-APR-2006 (first entry)
XX Tobacco cytochrome P450 cDNA clone D209-AA12, SEQ ID NO:220.
XX Plant breeding; crop improvement; secondary metabolite; genetic marker;
XX plant; cytochrome P450; CYP; gene; ss.
XX Nicotiana tabacum.
XX US2006041949-A1.
XX 23-FEB-2006.
XX 27-APR-2005; 2005US-00116881.
XX 13-NOV-2001; 2001US-0337684P.
XX 11-JAN-2002; 2002US-0347444P.
XX 12-MAR-2002; 2002US-0363684P.
XX 16-OCT-2002; 2002US-0418933P.
XX 13-NOV-2002; 2002US-00293252.
XX 10-JAN-2003; 2003US-00340861.
XX 12-MAR-2003; 2003US-00387346.
XX 08-JUL-2003; 2003US-0485368P.
XX 18-SEP-2003; 2003US-0503989P.
XX 16-OCT-2003; 2003US-00686947.
XX 29-APR-2004; 2004US-0566235P.
XX 03-SEP-2004; 2004US-00934944.
XX 03-SEP-2004; 2004US-0607357P.
XX 17-SEP-2004; 2004US-00943507.
XX 15-OCT-2004; 2004WO-US034065.
XX 15-OCT-2004; 2004WO-US034218.
XX 25-JAN-2005; 2005US-0646764P.
XX 24-MAR-2005; 2005US-0665097P.
XX 19-APR-2005; 2005US-0665451P.
XX 19-APR-2005; 2005US-00110062.
XX (USSM-) US SMOKELESS TOBACCO CO.
XX Xu D, Nielsen MT;
XX WPI; 2006-182895/19.
XX P-PSDB; AEF97227.
XX New breeding method, useful for producing a tobacco plant having
XX decreased expression of a nicotine demethylase gene comprises crossing
XX germinating F1 progeny seed to produce the tobacco plant.
XX Claim 63; SEQ ID NO 220; 51pp; English.
XX The invention relates to a breeding method for producing a tobacco plant
XX having decreased expression of a nicotine demethylase gene. The method
CC
```


CC involves crossing a first tobacco plant having variant nicotine
 CC demethylase gene expression with a second tobacco plant containing at
 CC least one phenotypic trait (e.g., disease resistance, high yield etc.) to
 CC produce an F1 progeny plant; collecting the seed of the F1 progeny; and
 CC germinating the seed to produce a tobacco plant having decreased nicotine
 CC demethylase expression. The invention also relates to a tobacco plant or
 CC its components produced using the method of the invention; a tissue
 CC culture of regenerable tobacco cells obtained from such plants; a tobacco
 CC product produced from such plants; and a method of breeding a tobacco
 CC plant with a modified attribute comprising variant expression of a
 CC cytochrome P450 polynucleotide (including the nicotine demethylase gene).
 CC The invention further relates to isolated constitutive, ethylene-induced
 CC or senescence-induced genetic markers comprising tobacco cytochrome P450
 CC nucleic acid molecules, including nicotine demethylase sequences; methods
 CC for reducing or increasing the expression or activity of polypeptides
 CC encoded by these nucleic acid molecules; expression vectors, plants or
 CC plant components comprising one of these polynucleotides; and tobacco
 CC products produced from such plants. The methods of the invention are
 CC useful for the breeding (especially marker assisted breeding) of tobacco
 CC plants with decreased nicotine demethylase expression or with altered
 CC cytochrome P450 expression or activity. Altered expression of such
 CC enzymes can result in a change in the composition of secondary
 CC metabolites such as alkaloids (e.g., nicotine), phenylpropanoids,
 CC terpenoids, lipids, cyanogenic glycosides and glucosinolates, with
 CC effects on the flavor or aroma of plant products. They may also affect
 CC herbicide tolerance, resistance to disease or insects, quality factors
 CC related to undesirable constituents, structural traits, fiber content,
 CC leaf yield, ripening, leaf curing or storage properties. Tobacco plants
 CC of the invention in which the expression of cytochrome P450 genes is
 CC altered may have desirable traits such as altered levels of nicotine
 CC or N'-nitrosomonocotine. Such plants can be used in the production of
 CC tobacco products such as moist or dry snuff, chewing tobacco, cigarettes,
 CC cigars, cigarillos, pipe tobacco, bidis or smokeless tobacco products.
 CC The present sequence represents a specifically claimed tobacco cytochrome
 CC P450 polynucleotide useful in marker assisted breeding methods. Note: The
 CC sequence data for this patent is also available in electronic format
 CC directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?DocID=20060041949.
 XX
 SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,41e-100 Length: 1673
 Score: 1149.00 Matches: 222
 Percent Similarity: 63.2% Conservative: 91
 Best Local Similarity: 44.8% Mismatches: 166
 Query Match: 44.6% Indels: 16
 DB: 15 Gaps: 5

US-10-759-813-2 (1-500) x AEF97226 (1-1673)

Qy 12 LeuileSerPheLeuLeuValLeuLeuValValValMetArgLeuTrpLysGln 31
 Db 54 TTGGTTTCATTTCTATTTCTCTTCCTCTTTTGTAAAGATATGGAAGAACTCC 113
 Qy 32 AsnPro-----ProGlyProTrpLysPheProIleileGlyAsnLeu 46
 Db 114 AATAGCCAAAGCAAAAGTTGCCACCAGGTCCATGGAACTACCAATACTAGGAAGTATG 173
 Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
 Db 174 CTTTCATATGGTT-----GGTGGACTACCAACACCACCATGCTCTTAGAGATTAGCCAAAAA 227
 Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerAla 86
 Db 228 TATGGACCACTTATGCACCTTCAATTAGGTGAAGTTTCGCGGTGTGGTTACTTCTCT 287
 Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
 Db 288 GATACGCGCAAGAAGTATTAAAACTCATGACATCGCTTTTGGGCTAGGCTAGCCTT 347
 Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
 Db 1428 AAACTCCCTGCTGGAATCGAACCAACGCGACTTGGAGCTTGAGTGTGGTAACT 1487

Db 348 TTGGCCCCGGAGATTCTCTGTGTACAAATAGGTCTGATCTAGCCTTTTGGCCCTATGGCGAC 407
 Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysValGln 146
 Db 408 TATTTGAGCAAAATGCGTAAATAATGTGCTTGGAAAGTGTCTAGTCCCAAGAAATGTTCGG 467
 Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
 Db 468 ACATTTAGTCTCTATTAGGCGGAATGAAGTTCTTCGTCTCATTAATTTTATCCGGTCTATCT 527
 Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
 Db 528 TCTGTGTAACCTTATTAACTTACGGAAGAGTCTTTTGTTCACAAGCTCCATGACATGT 587
 Qy 187 ArgThrSerValGlyAsnCys--LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
 Db 588 AGATCAGCGTTTGGCGMAGTGTTCAAAGAGCAAGCAAAATTTATACACTTAATTAAGAA 647
 Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
 Db 648 GTGATACTCTTAGCAGGAGGGTTTGATGTGGCTGACATATTCTCTTCACTGAAGTTTCTT 707
 Qy 226 HisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluLeuAspIle 245
 Db 708 CATGTCTCAGTGAATGAAGGTAAAGATTATGAATGCACACCATAGGTAGTGCATTT 767
 Qy 246 LeuGluGluLeuLeuAsnGluHis-----LysAlaLeuLysPro 258
 Db 768 GTTGAGATGTCTCAATGAGCACAAAGAAATCTTGGCAATTTGGGAAACTAATGAGCG 827
 Qy 259 PheGluAlaAsnLeuMetAspValLeuLeuLeuLeuLysAsnGlyAsnValPro 278
 Db 828 TTAGGAGGTGAAGATTATTAATTCATGTCTTCTTAAGACTTATGAATCATGAGGCGCTTCAA 887
 Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
 Db 888 TTTCTCTATCACCACGACCAATCAAGCCATAATTTTTTGACATGTTTCTGCGCGGACA 947
 Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
 Db 948 GAGACTTCTCTCAACAATTTGTGGGCTATGTTAGAAATGGTGAAAAATCCAGCGCTA 1007
 Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
 Db 1008 TTTCGCGAAAGCTCAAGCAGAGTGAAGAGAGCATTTAGAGGAAAAAGAAACTTTTCGATGAA 1067
 Qy 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
 Db 1068 AATGATGTGGAGGAGCTAAACTACCTAAAGTTAGTAATAAAAAAGAACTCTAAGACTTCAT 1127
 Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
 Db 1128 CCACCGGTTCCACTTTTGTCTCCAGAGAGATGTAGGGAAGAGACAAATATAACGGCTAC 1187
 Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
 Db 1188 ACTATTCTCTGTAAGACCAAAAGTCATGTTTATGTTGGGCTTTGGGAAGAGATCCAAAA 1247
 Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
 Db 1248 TATTGGAATGACGCAAGAACTTTTATGCCAGAGAGATTGTAGCAGTGTCTCAAGGATTTT 1307
 Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
 Db 1308 GTTGGTAATAATTTTGAATATCTTCATTTTGTGGCGGAAGAGAGATTTGCTCTGGATT 1367
 Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleAsnLeuLeuTyrHisPheAsnTrp 457
 Db 1368 TCGTTTGGCTTAGCTAAATGCTTATTGTCATTTGGCTCAATTACTATATCATTCTGATTGG 1427
 Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
 Db 1428 AAACTCCCTGCTGGAATCGAACCAACGCGACTTGGAGCTTGAGTGTGGTAACT 1487

Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProileProTyrGln 492
Db 1488 GCCGCTAGAGAAAAAGTGACCTTTACTTGGTTGGGACTCCTTATCAA 1532

Search completed: May 31, 2006, 06:42:14
Job time : 939 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:42:32 ; Search time 1730 Seconds
(without alignments)
5327.010 Million cell updates/sec

Title: US-10-759-813-2
Perfect score: 2574
Sequence: 1 MBQKNLSPSILISFLVLI.....KIDKLPIPIYQVSLGNSIS 500

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n model -DEV=xlh
-Q=abss/ABSSWEB_spool/US10759813/runat_31052006_060313_2784/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss05h
-USER=US10759813 @CGN_1_1675@runat_31052006_060313_2784 -NCPU=6 -ICPU=3
-NO_WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :				Published Applications NA Main:			
1:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:						
2:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:						
3:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:						
4:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:						
5:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:						
6:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:						
7:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:						
8:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:						
9:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:						
10:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:						
11:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:						
12:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:						
13:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:						
14:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:						
15:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:						
16:	/EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2574	100.0	1733	3	US-09-909-566C-1

2	2574	100.0	1733	8	US-10-759-813-1	Sequence 1, Appli
3	1164	45.2	1538	8	US-10-686-947-153	Sequence 153, App
4	1164	45.2	1538	10	US-10-943-507-153	Sequence 153, App
5	1164	45.2	1538	11	US-10-934-944-153	Sequence 153, App
6	1164	45.2	1538	16	US-11-116-881A-162	Sequence 211, App
7	1149	44.6	1673	8	US-10-686-947-211	Sequence 231, App
8	1149	44.6	1673	8	US-10-686-947-231	Sequence 208, App
9	1149	44.6	1673	11	US-10-943-507-208	Sequence 211, App
10	1149	44.6	1673	11	US-10-934-944-211	Sequence 231, App
11	1149	44.6	1673	11	US-10-934-944-231	Sequence 220, App
12	1149	44.6	1673	16	US-11-116-881A-220	Sequence 240, App
13	1149	44.6	1673	16	US-11-116-881A-240	Sequence 209, App
14	1148	44.6	1673	10	US-10-943-507-209	Sequence 206, App
15	1148	44.6	1673	10	US-10-943-507-206	Sequence 209, App
16	1148	44.6	1673	11	US-10-934-944-209	Sequence 218, App
17	1148	44.6	1673	16	US-11-116-881A-218	Sequence 195, App
18	1144	44.4	1657	8	US-10-686-947-195	Sequence 192, App
19	1144	44.4	1657	10	US-10-943-507-192	Sequence 195, App
20	1144	44.4	1657	11	US-10-934-944-195	Sequence 204, App
21	1144	44.4	1657	16	US-11-116-881A-204	Sequence 213, App
22	1142	44.4	1673	8	US-10-686-947-213	Sequence 210, App
23	1142	44.4	1673	10	US-10-943-507-210	Sequence 213, App
24	1142	44.4	1673	11	US-10-934-944-213	Sequence 197, App
25	1142	44.4	1673	16	US-11-116-881A-222	Sequence 261, App
26	1131	43.9	1610	8	US-10-686-947-197	Sequence 194, App
27	1131	43.9	1610	8	US-10-686-947-261	Sequence 256, App
28	1131	43.9	1610	10	US-10-943-507-194	Sequence 197, App
29	1131	43.9	1610	10	US-10-943-507-256	Sequence 261, App
30	1131	43.9	1610	11	US-10-934-944-261	Sequence 206, App
31	1131	43.9	1610	16	US-11-116-881A-206	Sequence 270, App
32	1131	43.9	1610	16	US-11-116-881A-270	Sequence 149, App
33	1131	43.9	1610	8	US-10-686-947-149	Sequence 149, App
34	1129.5	43.9	1576	10	US-10-943-507-149	Sequence 158, App
35	1129.5	43.9	1576	11	US-10-934-944-149	Sequence 191, App
36	1129.5	43.9	1576	16	US-11-116-881A-158	Sequence 188, App
37	1129.5	43.9	1576	8	US-10-686-947-191	Sequence 191, App
38	1128.5	43.8	1566	10	US-10-943-507-188	Sequence 191, App
39	1128.5	43.8	1566	11	US-10-934-944-191	Sequence 200, App
40	1128.5	43.8	1566	16	US-11-116-881A-200	Sequence 151, App
41	1128.5	43.8	1581	8	US-10-686-947-151	Sequence 151, App
42	1128.5	43.8	1581	11	US-10-943-507-151	Sequence 160, App
43	1128.5	43.8	1581	11	US-10-934-944-151	
44	1128.5	43.8	1581	16	US-11-116-881A-160	
45	1128.5	43.8	1581	16	US-11-116-881A-160	

ALIGNMENTS

RESULT 1
US-09-909-566C-1
; Sequence 1, Application US/09909566C
; Publication No. US20030066103A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-
; FILE OF INVENTION: 13-epoxy fatty acids
; FILE REFERENCE: BBI465 US NA
; CURRENT APPLICATION NUMBER: US/09/909,566C
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219833
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1733
; TYPE: DNA
; ORGANISM: Euphorbia lagascae
US-09-909-566C-1

Alignment Scores:
Pred. No.: 6.66e-282
Score: 2574.00
Length: 1733
Matches: 500
Percent Similarity: 100.0%
Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-759-813-2 (1-500) x US-09-909-566C-1 (1-1733)

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
DB 14 ATGGAGCAGAAAAATCTCTCTTTCCGAGCATTTTAAATAAGTTTCTGCTGTTTTAATC 73

QY 21 LeuValValMetArgLeuTrpLysLysGlnAsnProProGlyProTrpLysPhe 40
DB 74 TTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGGGCCATGGGAAGTTT 133

QY 41 ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe 60
DB 134 CCTATCATAGGTAAATCTCTCTTATTAATCTCACTTCTGATCTAGGCCATGAAGCTTTT 193

QY 61 ArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
DB 194 AGAGGCTTTGGCTCAAAATTTATGGAGCTGTATGAGTCTTCAAAATTTGGCCAAGTTTCAGCT 253

QY 81 ValValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPhe 100
DB 254 GTTGTCATTTCTTTCAGCTGAAGCAGCAAGAGGTTATGAAGCTCAGGCTGATGCCCTTC 313

QY 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu 120
DB 314 GCCCAACGCCCTATCGTCTGGACGCACAGATTGTGTTTATAATCGGAAGATGCTCTG 373

QY 121 PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
DB 374 TTTGCTTCATATGGAGATCACTGGAGGCAGATGAAGAAAAATTTGGATCTACTGAAATTTCTG 433

QY 141 SerAlaLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIle 160
DB 434 AGTGCAAAAAAGTTCAATCTCTCCAGGTTAATCCGAGAGGAAGAAATGGAGGATCCCATC 493

QY 161 ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIle 180
DB 494 ACATTTCTCCGTTCCGAAGCCGATCTCCGGTCAATATTACAAAGATCATTTATGGCAT 553

QY 181 IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu 200
DB 554 ATAATTTCCGATCATGATGAAGACATCCGTTGGTAATTTGAAGCAAAAAAGAAATGCTG 613

QY 201 SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro 220
DB 614 AGTGTTCGCGATGCAGTCNATGAGCGCGCAGGATTTTGGCACCAGCAGCGCTTTTCG 673

QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgLeuHisGln 240
DB 674 ACGTGGAAATTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAGCGGTTTGCATCAG 733

QY 241 GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLysProPheGlu 260
DB 734 GAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCCAATGAAGCTTTTGA 793

QY 261 AlaAspAsnLeuMetAspValLeuLeuLeuAsnGlnLysAsnGlyAsnValProValPro 280
DB 794 GCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTCCAGTGCCA 853

QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
DB 854 GTGCAAAACGAAGCATCAAGCATCCGTTTGTGCAAAATGTTTACTGCGGGAGCGCAACA 913

QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
DB 914 ACTTCGAAGCTPACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCCAACTGAATGA 973

QY 321 LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
DB 974 AAAGCACAAGAGAAGTTAGACAAGTATTTTGTGAAATGGGAAAAAGTTGATGAATCAAGA 1033

QY 341 PheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisProPro 360
DB 1034 TTTTCATGATTTGAAATTTCTTCAAGTTAGTGTGTTAAAGAAACTCTAAGATTACATCTCCG 1093

QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
DB 1094 GTTGTCTTGATTTCCGAGGGAGGTAGAGAAACAAACCAATTTGATGGGATATGAAATTCAT 1153

QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
DB 1154 CCGAACACTCGAATTTGTGATGCTTGGGGGATAGGAAGATCTCTAATACTTTGGTCG 1213

QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
DB 1214 GAACCTGGAAAGTTTAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATATAAAGGAGC 1273

QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
DB 1274 ACATTTGAACCTGGTACCATTTGGTGCAGGAAAAAGAAATATGTCTGGCATTTACTTCAGCT 1333

QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
DB 1334 ATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAAATGGGAACCTGCC 1393

QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
DB 1394 GATGGAATTCACCTCAAAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAA 1453

QY 481 LysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeuGlySerAsnIleSer 500
DB 1454 AAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTCAAAATATTCT 1513

RESULT 2

US-10-759-813-1

; Sequence 1, Application US/10759813

; Publication No. US20040139499A1

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Edgar B

; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-

; FILE REFERENCE: BBI465 US NA

; CURRENT APPLICATION NUMBER: US/10/759,813

; CURRENT FILING DATE: 2004-01-15

; PRIOR APPLICATION NUMBER: US/09/909,566C

; PRIOR FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: 60/219833

; PRIOR FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1733

; TYPE: DNA

; ORGANISM: Euphorbia lagascae

US-10-759-813-1

Alignment Scores:

Pred. No.: 6,66e-282 Length: 1733

Score: 2574.00 Matches: 500

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 8 Gaps: 0

US-10-759-813-2 (1-500) x US-10-759-813-1 (1-1733)

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
DB 14 ATGGAGCAGAAAAATCTCTCTTTCCGAGCATTTTAAATAAGTTTCTGCTGTTTTAATC 73

QY 21 LeuValValMetArgLeuTrpLysLysGlnAsnProProGlyProTrpLysPhe 40
DB 74 TTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGGGCCATGGGAAGTTT 133

```
QY 41 ProIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe 60
DB 134 CCTATCATAGTAATCTTCTCATTTATTACTACATCTTGATCTAGGCCATGAACGTTTT 193
QY 61 ArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
DB 194 AGAGCCCTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGGCCCAAGTTTCAGCT 253
QY 81 ValValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaSerAlaPhe 100
DB 254 GTTGTCATTTCTTCACTGAAGCAGCAAGAGGTTATGAAAACTCAGGCTGATGCCCTTC 313
QY 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu 120
DB 314 GCCCAACGCCCTATCGCTTGGACGCGACAGATTTGTTTTATTAATCGGAAGATGCTTGG 373
QY 121 PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
DB 374 TTTGCTTCATATGGAGATCACTGGAGCGCAGATGAAGAAAAATTTGGATACTTGAATTTCTG 433
QY 141 SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIle 160
DB 434 AGTGCCAAAAAAGTTCAATCTCCAGGTTAAATCCGAGAGGAAGAAATGGAGGATGCCATC 493
QY 161 ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIle 180
DB 494 ACATTCCTCGTTCGNAACCGGATCTCCGGTCAATATTACAAGATCATTTATGGCAT 553
QY 181 IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu 200
DB 554 ATAATTTCCATCATGATAAGAACATCCGTTTGGTAAATTTGAAGCAAAAGAAAGATTGCTG 613
QY 201 SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaSerAlaPhePro 220
DB 614 AGTGTCGCGATCGAGTCAATGAGCGAGCAGATTTTGGCACCACCGCATTTTTCG 673
QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
DB 674 ACGTGGAAATTACTTCACTATATCATTTGCGAGCTGAGTCAAAACCCAGCGGTTTGCATCAG 733
QY 241 GluIleAspAspIleLeuGluLeuLeuAsnGluHisLysAlaAsnLysProPheGlu 260
DB 734 GAGATTGACGATATCTTGAAGAGATTCTTAATGAACAAAGCCCAATAAGCCTTTTGA 793
QY 261 AlaAspAsnLeuMetAspValLeuLeuLeuGlnLysAsnGlyAsnValProValPro 280
DB 794 GCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTCCAGTGCCA 853
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
DB 854 GTGACAAACGNAAGCATCAAGCATCCGTTTGGCAAAATGTTTACTGCCGGGAGCGNAACA 913
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
DB 914 ACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGA 973
QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
DB 974 AAAGCACAGAAGAGATTAGACAAGTATTGTGGAATGGGAAAAATGTAATGAATCAAGA 1033
QY 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
DB 1034 TTTTCATGATTGAAATCTTCAAGTAGTGGTTAAAGAAACTCTAAGATTATCATCTCTCG 1093
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluLeHis 380
DB 1094 GTTGCTTGTATTCGAGGAGGTGTAGAGAAACAACACAGAAATGATGATGAATGAATTCAT 1153
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
DB 1154 CCGAACACTCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCCCTAATACTTGGTGC 1213
```

```
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
DB 1214 GAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTGATTATAAAGGACG 1273
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
DB 1274 ACATTTGAACCTGGTACCATTTGGTGCAGGAAAAAANAATATGTCTCGCATTTACTCAGCT 1333
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
DB 1334 ATTACCAATTTGAGATGATGTCATTATAAATCTATTATATCATTTTAATTTGGGAACGTGCC 1393
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
DB 1394 GATGGAATTTACACCTCAACACCTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAA 1453
QY 481 LysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeuGlySerAsnIleSer 500
DB 1454 AAAATAGATCTTAAAGTTGATTCTCTATTCCATATCAAGTTAGCTTAGCTCAAAATATTCT 1513

RESULT 3
US-10-686-947-153
; Sequence 153, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-153

Alignment Scores:
Pred. No.: 2,77e-121 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
DB: 8 Gaps: 5

US-10-759-813-2 (1-500) x US-10-686-947-153 (1-1538)
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysGln 31
DB 48 TTGGTTTCCATTTTCTTCTATTCTTTCTTTTCTTTTAAAGAAATGGAAGAACTCG 107
QY 32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
DB 108 ATATGCCAAAGCAAAATAATTCGCCACCGTCCATGGAAACTACCAATACATGAGAGTATG 167
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
DB 168 CTTTCATATGGTT-----GGTGGACTACCAACCATGTCCTTTAGAGATTTAGCCAAAAA 221
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
DB 222 TATGACCGCTTATGACCTTCATAGGTGAAGTTTCTGCAGTTGTGTTTCTTCTCCT 281
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
DB 282 GATATGGCAAAAGAAAGTACTAAAAACTCATGACATCGCTTTTCGGTCTAGGCCCTT 341
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
DB 342 TTGGCCCGGAGATTGTTCTGTACAAATAGTCTGATCTTTGCGCTTTTGGCCCTATGCGCAT 401
```

```
Qy 127 HistPrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGGAGACAAATCGTAAATAATGTGCTTGGAGTGCCTCAGTGCACAAGAAATGTTCCG 461
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 462 ACATATAGCTCTATTAGCGCATGAAGTCTCTCGTCTCTTAATTTATTCGGGTCATCT 521
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTrpGlyIleIleSerIleMetIle 186
Db 522 TCTGTGAGCCCTGTAATATTACGAAGAGATCTTTTGTTCACAAAGCTCATGACATGT 581
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuSerValAlaAspAla 205
Db 582 AGATCAGCGTTTGGGCAAGTATTCAGGAGCAAGCAAAATTTATACAACATAATTAAGAA 641
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 642 GTTATACTCTTAGCAGGAGGGTTTCATGTGGCTGCATATTATCCCTTCATACAAGTCTCT 701
Qy 226 HistTrpIleIleGlyAlaGlySerLysProArgLeuHisGlnGluIleAspIle 245
Db 702 CATGTGCTCAGTGAATGAGGTAAGATTATGAATGCACACCATTAAGTAGATCTATT 761
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 762 GTTGAGATGTCATCAACGAGCACAAAGAAATCTTGCAATTTGGGAAACATAATGGAGCG 821
Qy 259 PheGluAlaAspAsnMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 822 TTAGAGGTGAAGATTAAATGATGTTCTTCTAAAACCTATGAATGATGGAGCGCTTCAA 881
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 882 TTTCTATCACCAACGACACATCAAGCTATAATCTTTGACATGTTGCTGCTGGAACA 941
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 942 GAGACTTCATCGTCAACAATTTGTGGGCTATGGTGGAAATGGTGAAATTCCTCAACTGTA 1001
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1002 TTTGCGAAGCTCAACGAGAGTAAAGATGATTCATTAGAGAAAGAAACATTTTGATGAA 1061
Qy 339 SerArgPheHisAspLysPheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1062 AATGATGTGGAGAGCTAAACTATCTAAAGTTAGTCAATTAAGAAACTCTAAGACTTCAT 1121
Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1122 CCACCGGTTCCACTTTGCTCCAAAGAGAATGTAGGGAAGACAAATATTAACCGCTAC 1181
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTtpAlaIleGlyArgAspProAsn 397
Db 1182 ACTATTCTGTAAAGACCAAGATCATGTTAATGTTTGGGCATTGTGGAGAGATCCAAA 1241
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1242 TATTGGGATGATGCAGAACTTTTAAGCCAGAGAGATTTGACGAGTGCCTCAAGGATTT 1301
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1302 GTTGGTAATAATTTTGAATATCTTCATTTGGTGGTGGAGGAGGATTTGTCAGGGATT 1361
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1362 TCGTTTGGTTTAGCTAATGCTTATTGTCATTTGGCTCAATTACTTTATCACTTTGATTGG 1421
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1422 GAACCTCCCACTGGAATCAACCAACGAGCTTTGGACTTGACTGAGTTGGTTGGAGTAAC 1481
```

```
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1482 GCCGCTAGAAAAGTGACCTTTACTTGGTTGGACTCCTTATCAA 1526

RESULT 4
US-10-943-507-153
; Sequence 153, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-943-507-153

Alignment Scores:
Pred. No.: 2,77e-121 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
Gaps: 5

US-10-759-813-2 (1-500) x US-10-943-507-153 (1-1538)
Qy 12 LeuLeSerPheLeuLeuValLeuValValMetArgLeuTrpLysLysGln 31
Db 48 TTGGTTTCCATTTTCTATTTCTTTCTTTTCTTTTAAAGAAATGGAAGAACTCG 107
Qy 32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 108 AATAGCCAAAGAAAAAATTCACCAGTCCATGGAAACTTACCANTACTAGGAAGATG 167
Qy 47 ProHisLeuLeuLeuThrSerAspLysGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 168 CTTCATATGGTT-----GGTGGACTACCACCATGCTCTTAGAGATTTAGCCAAAAA 221
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
Db 222 TATGACCGCTTATGACCTTCAATTAGGTGAAAGTTTCTGCAAGTTGTGTTACTTCTCT 281
Qy 87 GluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 282 GATATGGCAAAAGAGTACTAAAAACTCATGACATCGCTTTTCGCCTCTAGGCCTAGCCCT 341
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 342 TTGGCCCGGAGATTGCTGTACAAATAGGTCTGATCTTGGCTTTTGGCCCTATGGCGAT 401
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGGACAAATGCGTAAAAATATGTCTTCTTGGAAAGTCTCAGTGCACCAAGAAATGTCG 461
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
```

```

...
462 ACATATAGCTCTATTAGGCGCGATGAAGTTCTTCGTCTCTAAATTTTATCCGGTCATCT 521
QY
167 AlaGlySerProValAsnIleThrLysIleIleThrGlyIleIleSerIleMetIle 186
...
522 TCTGGTACCGCTGTTAATATTACGAAAGAGATCTTTTGTTCACAGCTCCATGACATGT 581
QY
187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
582 AGATCAGCGTTTGGGCAAGTATTCAAGGACGACAAATTTATACAACCTAATTAAGAA 641
Db
206 ValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
642 GTTATACTCTTAGCAGGAGGGTTTGATGTGGCTGACATATTCCCTTCATACAAGTCTCT 701
QY
226 HisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspIle 245
702 CATGTCTCAGTGAATGAAGGGTAAGATTATGAATGCACCATTAAGGTAGTACTATT 761
QY
246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
762 GTTGAGAATGTCATCAACGACGACAGAAAATCTTGCAATTGGGAAAACCTAATGAGCG 821
Db
259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
822 TTAGGAGGTGAAGATTAAATGATGTCTCTAAAACCTTATGAATGATGAGCGCTTCAA 881
QY
279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
882 TTTCTCTATCACCACGACACATCAAAAGCTATATCTTTGACATGTTTCTGCTGGAA 941
QY
299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
942 GAGACTTCATCGTCAACATTTGTGGGCTATGGTGGAAATGGTGAAAATCCAACTGTA 1001
Db
319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
1002 TTTGCGAAAGCTCAAGCAGACAGTAAGAGATGCATTTAGAGAAAAGAAAACCTTTGATGAA 1061
QY
339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
1062 AATGATGTGGAGAGCTAAACTATCTAAAGTTAGTCAATTAAGAAGAACTCTAAGACTTCAT 1121
QY
359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
1122 CCACCGTTCCACTTTGCTCCAGAGAAATGTAGGGAAGACAAATATAACGGCTAC 1181
QY
378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
1182 ACTATTCTCTGTAAGACCAAGATCATGTTGTTGGCATTTAGGCAAGAGATCCAAA 1241
QY
398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
1242 TATTGGGATGATGCGAAACTTTTAAAGCCAGAGAGATTTGAGCGTCTTAAGGATTTT 1301
Db
418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
1302 GTTGTAAATATTTGAATATCTTCCATTTGTTGGTGGAGGAGGATTTGTCAGGGATT 1361
QY
438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
1362 TCGTTTGGTTAGCTAATGCTTATTGTCATTTGGCTCAATTACTTTATCACTTGATTGG 1421
QY
458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
1422 GAACATCCCACTGGAATCAACCAACGACCTTGGACTTGGAGTGGTGGAGTAACT 1481
QY
478 LeuArgLysIleAspLeuLysLeuIleProIleProTyrGln 492
1482 GCGCTAGAAAAGTGACCTTTACTTGGTGGACCTCTTATCAA 1526

```

RESULT 5

US-10-934-944-153

```

; Sequence 153, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; US-10-934-944-153

Alignment Scores:
Pred. No.: 2,77e-121 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
DB: 11 Gaps: 5

US-10-759-813-2 (1-500) x US-10-934-944-153 (1-1538)

QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db 48 TTGGTTTCCATTTTCCATATTTCTATCTTTCTTTTAAAGGAAATGGAAGAACTCG 107
QY 32 AnPro-----ProProGlyProTrpLysPheProIleLeuGlyAsnLeu 46
Db 108 AATAGCCAAAGGAAAAAATTGCCACCATGCTCCATGAAACTACCAATATCTAGGAAGATG 167
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 168 CTTTCATATGGTT-----GGTGGACTACCAACCATGCTCTTAGAGATTTAGCCAAAAA 221
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 222 TATGACCGCTTATGCACCTTCAATTAGGTGAAGTTCTCTGCAAGTTGTGGTTACTTCTCCT 281
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 282 GATATGGCAAAAGAAAGTAGTACAAAAAATCATGACATCGCTTTTCGCGCTTAGGCCTTAGCCTT 341
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 342 TTGGCCCGGAGATTGTCTGTGTACATAGGTCTGATCTTGGCTTTGGCCCTATGGCGAT 401
QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGGAGACAAATGCGTAAAAATATGTCTTGGAAAGTGCTCAGTGTCCAAGAAATGTTGG 461

```



```
QY 206 ValAenGluAlaThrSerPheGlyThrAlaAaspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACCTCTTAGCAGGAGGGTTTGTATGTGGCTGACATATTCCTTCTACTGAAGTTTCTT
QY 226 HistyrllelleGlyAlaGluSerLysProArgArgLeuHisGlnGluLeuAaspAspIle 245
Db 708 CATGTGCTCAGTGGGAATGAAGGGTAAGATTATGAATGCACACCATTAAGGTAGATGCCATT 767
QY 246 LeuGluGluIleLeuAenGluHis-----LysAlaAenLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAGAAAATCTTGCAATTTGGGAAAACAACTAATGGAGCG 827
QY 259 PheGluAlaAaspAenLeuMetAaspValLeuLeuAenLysGlnLysAenGlyAenValPro 278
Db 828 TTAGGAGGTGAAGATTAAATGATGTTCTTCTTAAGACTTATGAATGATGAGGCGCTTCAA 887
QY 279 ValProValThrAenGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCCTATCACCACCAACATCAAGCCATAATTTTGTGACATGTTTGTGCGCGGACA 947
QY 299 GluThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAenProThrGlu 318
Db 948 GAGACTTCATCTCAACAATTTGTGGCTATGTTAGTAAATGGTGAAAATCCAGCCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAaspGlu 338
Db 1008 TTCCGCAAGCTCAAGCAGAAAGTAAGAGAACATTTAGAGGAAAAGAACTTTTCGATGAA 1067
QY 339 SerArgPheHisAaspLeuLysPheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGCTAACTACCTAAAGTTAGTAAATAAAGAACTCTAAGACTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAaspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTGCTCCAGAGAAATGTAGGGAAGACAAATAATAACCGCTAC 1187
QY 378 GluIleHisProAenThrArgIleValValAenAlaTrpAlaIleGlyArgAaspProAen 397
Db 1188 ACTATTCTCTGTAAGACCAAAAGTCATGTTTAAATGTTTGGGCTTTGGGGAAGAGATCCAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAenProGluArgPheLysAaspCysAlaIleAaspTyr 417
Db 1248 TATTGGAATGACGCAAACTTTTATGCCAGAGATTTTGACAGTGTCTTAAGGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGGTAATAATTTTGAATATCTTCCATTTGTGGCGGAAGAGGATTTGCTCGGATT 1367
QY 438 ThrSerAlaIleThrAenLeuGluTyrValIleIleAenLeuLeuTyrHisPheAenTrp 457
Db 1368 TCGTTTGGCTTAGCTAATGCTTATTGTCATTTGGCTCAATTACTATATCATTCACTTCGATTGG 1427
QY 458 GluLeuAlaAaspGlyIleThrProGlnThrLeuAaspMetThrGluAlaIleGlyAla 477
Db 1428 AAATCCCTGCTGGAAATCGAACCAAGCGACTTGGACTTGACTGAGTGGTGGAGTAAC 1487
QY 478 LeuArgLysLysIleAaspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTTGGTTGGGACTCTCTTATCAA 1532
```

RESULT 8

```
US-10-686-947-231
; Sequence 231, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
```

```
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 231
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-231
```

Alignment Scores:

Pred. No.:	1.62e-119	Length:	1673
Score:	1149.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	44.8%	Indels:	16
DB:	8	Gaps:	5

US-10-759-813-2 (1-500) x US-10-686-947-231 (1-1673)

```
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGGTTTCCATTTTCCTTATTTCTTCTTCTTTTGTAGGATATGAGAACTCC 113
QY 32 AenPro-----ProProGlyProTrpLysPheProIleIleGlyAenLeu 46
Db 114 AATAGCCAAACAAAAGTTGCCACCAAGTCCATGGAAACTACCAATACTAGGAAGTATG 173
QY 47 ProHisLeuLeuLeuThrSerAaspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCATATGGTT-----GGTGGACTACCAACACCATGTCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGACCACTTATGCACTTCAATTAGTGGAAGTTTCTGCGTTTGGTTACTTCTCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAaspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGCCAAAGAAGTATTAAAACTCATGACATCGCTTTTGGCTTAGGCCTAGCCTT 347
QY 107 LeuAaspAlaGlnIleValPheTyrAenArgLysAaspValLeuPheAlaSerTyrGlyAasp 126
Db 348 TTGGCCCGGAGATTGCTGTACAAATAGGTCTGTATCTAGGCTTTTGGCCCTTATGCCGAC 407
QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATGCGTAAATATGTCTTGGAAAGTCTCAGTCCCAAGAAATGTCGG 467
QY 147 SerSerArgLeuIleArgGluGluMetGluAaspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTCTGTCTCAATTAATTTATCCGTCATCT 527
QY 167 AlaGlySerProValAenIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGGTGAACCTATTAAATGTTACGGAAGAGGATCTTTTGTTCACAAGTCCATGACATGT 587
QY 187 ArgThrSerValGlyAenCys---LysGlnLysGluArgLeuLeuSerValAlaAaspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCNAAGAGCAAGCAAAATTTATACAACTAATTAAGAA 647
QY 206 ValAenGluAlaIaThrSerPheGlyThrAlaAaspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTTAGCAGGAGGGTTTGTATGTGCTGACATATTCCTTCTACTGAAGTTTCTT 707
QY 226 HistyrllelleGlyAlaGluSerLysProArgArgLeuHisGlnGluLeuAaspAspIle 245
Db 708 CATGTGCTCAGTGGAAATGAAGGTAAGATTATGAATGCACACCATTAAGGTAGATGCCATT 767
QY 246 LeuGluGluIleLeuAenGluHis-----LysAlaAenLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAGAAAATCTTGCAATTTGGGAAAACAACTAATGGAGCG 827
QY 259 PheGluAlaAaspAenLeuMetAaspValLeuLeuAenLysGlnLysAenGlyAenValPro 278
Db 828 TTAGGAGGTGAAGATTAAATGATGTTCTTCTTAAGACTTATGAATGATGAGGCGCTTCAA 887
```

```
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTCTATCACCACACACATCAAGCCATATTTTTCACATGTTTGGTCGGGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTCATCGTCAACAATTTGTGGCTGTATGTTAGTAATGTTGAAATATCGTCAAGCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCCGAAGAGCTCAAGCAGAAGTAAGAGAGCAATTTAGAGAAAAGAACTTTCCGATGAA 1067
QY 339 SerArgPheHisAspLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGCTAACTAACTAAAGTTAGTAATAAAGAACTCTAAGACTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgLysAspGlyTyr 377
Db 1128 CCACCGGTTTCCACTTTTGTCTCCCAAGAGATGTAGGGAAGAGACAAATATAAACGGCTAC 1187
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCTCTGTAAGACCAAGTCATGGTTAATGTTGGGCTTTGGGAAGAGATCCAAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAAACTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgLysCysProGlyIle 437
Db 1308 GTTGGTAATAATTTTGAATATCTCCATTTGGTGGCGAAGGAGGATTTGCTCGGATT 1367
QY 438 ThrSerAlaIleThrAsnLeuGluTrpValIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTGGCTTAGCTAATGCTTATTTGCCATTGGCTCAATTAATTAATCACTTCGATTGG 1427
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AACTCTCTGCTGGATCGAATCGAACCAAGCGACTTGGACTTGCAGTGTTGGTGAAGTAACT 1487
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCGGCTAGAAAAAGTGACCTTTACTTGGTTGGGACTCCTTATCAA 1532

RESULT 9
US-10-943-507-208
; Sequence 208, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 208
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
```

US-10-943-507-208

```
Alignment Scores:
Pred. No.: 1,62e-119 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.8% Indels: 16
DB: 10 Gaps: 5
```

US-10-759-813-2 (1-500) x US-10-943-507-208 (1-1673)

```
QY 12 LeuIleSerPheLeuLeuValValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGGTTTCCATTTCTCTATTTCTATCTTTCTCTCTTTTGTAGGATATAGGAAGAACTCC 113
QY 32 AsnPro-----ProProGlyProTrpLysPheProIleLeuGlyAsnLeu 46
Db 114 AATAGCCAAAGCAAAAGTTGCCACCAGGTCCATGGAAACTACCAATACTAGGAAGTATG 173
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCTATGCTGTT-----GGTGGACTACCAACCATGTCTCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGGACCACTTATGACCTTCATATTAGGTGAAGTTTCTGCGTTGTGGTTACTTCTCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGCCAAAGAGATTAATAAACTCATGACATCGCTTTTGGCGTCTAGGCTTAGGCTT 347
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCCGGAGATTGTCTGTACAAATAGGTCTGTAGCTCTTGGCCCTATGCGGAC 407
QY 127 HistTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATCGGTAAATATGCTGCTTGGGAAGTCTCAGTGCACCAAGATGTTCCG 467
QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGGCGGAATGAAGTTCTTCGTCTCATTAATTTATTCGGTCTATCT 527
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
Db 528 TCTGTGAACCTTATTAATGTTTACGGAAGGATCTTTTGTTCACAAGCTCCATGATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGGCAAGTGTTCAAAGAGCAAGCAACNAATTTATACAACCTAATTAAGAA 647
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGTTTGTATGTGGCTGACATATTCCTTCACTGAAGTTTCTT 707
QY 226 HistTyrIleIleGlyAlaGluSerLysProArgLysGlnLeuHisGlnLeuLeuAspIle 245
Db 708 CATGTGTCAGTGAATGAAGGTAAGATTATGAATGTCACACCACTAAGGTAGTAGCATT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAGAAAATCTTGCATTTGGGNAAACTAATGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTTAATGATGTTCTTCTTAAGACTTATGAATGATGAGGCGCTTCAA 887
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTCTATCACCAACGACCAACATCAAGCCATAATTTTGGACATGTTTGTCTCCCGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
```

```
Db 948 GAGACTTCATCGTCACCAATGTGTGGCGCTATGGTAGAAATGGTGAAGAAATCCAGCCGCTA 1007
Qy 319 LeuArgLysAlaGlnGluGlnValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGGCAAGCTCAGACGAGAGTAGAGAGCAATTTAGAGGAAAGAAACTTTTCGATGAA 1067
Qy 339 SerArgPheHisAspLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGCTAACTACCTACCTAAAGTTAGTAAATAAAGAAACTCTAAGACTTCAT 1127
Qy 359 ProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTTGTCTCCAGAGAAATGTAGGGAAGACAAATATAAACGGCTAC 1187
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTtpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCCTGTAAAGCAACCAAGTCATGTTAATGTTTGGCTTTGGGAAGAGATCCAAA 1247
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACCGCAAACTTTTATGTCAGAGAGATTGAGCAGTGCTCTAAGGATTTT 1307
Qy 418 LysGlyThrThrPheGlyLeuValProPheGlyAlaGlyLysArgIleCysProGlyTle 437
Db 1308 GTTGTGTAATAATTTTGAATATCTTCATTTGGTGGCGGAAGGAGATTGTCTCGGATT 1367
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTAGCTAATGCTTATTGGCAATTGGCTCAATTACTATATACATTCGATTGG 1427
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1428 AAACTCCCTGCTGGAATCGAACCAAGCGACTTTGGACTTGCTGAGTGTGTTGGAGTAAC 1487
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGCACTCTTATCAA 1532
```

RESULT 10

```
US-10-934-944-211
; Sequence 211, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 211
; LENGTH: 1673
```

```
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-934-944-211
Alignment Scores:
Pred. No.: 1,62e-119 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
Gaps: 5
DB:
US-10-759-813-2 (1-500) x US-10-934-944-211 (1-1673)
Qy 12 LeuIleSerPheLeuLeuValValLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGGTTTCCATTTTCTTATCTTCTCTCTTTTGTAAAGGATATGAAGAACAATCC 113
Qy 32 AsnPro-----ProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 AATAGCCAAAGCAAAAAGTTGCCACAGGTCCATGGAAACTACCAATATCTAGGAAGTAG 173
Qy 47 ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheAlaLeuAlaGlnIle 66
Db 174 CTTCATATGGTT-----GGTGGACTACCAACCATGCTCTTAGAGATTTAGCCAAAAA 227
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyValSerAlaValValIleSerSerAla 86
Db 228 TATGACCCACTTATGCACTTCAATTAGGTGAAGTTCTCGCGTTGTGGTTACTTCTCT 287
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAAAGAGATTTAAAACTCATGACATCGCTTTTGGCTCTAGCCCTAGCCTT 347
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGCCCCGGAGATTGCTGTACAAATAGGTCTGATCTAGCCTTTTGGCCCTATGCGCAC 407
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATGCGTAAATATGTCTTGGAAAGTGCCTGAGTGCAGCAAGATTTCCG 467
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTCTGCTCATTAATTTTATCCGGTCACT 527
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGTTGAACCTATTATATGTTACGMAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCAAAGAGCAAGACAAATTTATACAACCTAATTAAGAA 647
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATCTCTTAGCAGGAGGGTTTGATGTGCTGCATATTCCTTCTCAAGAGTTCTT 707
Qy 226 HistTrpIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspIle 245
Db 708 CATGTGCTCAGTGGAAATCAAGGTAAGATTATGAATGCACACCAATAAGTAGATGCCATT 767
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGAAATGTCATCAATGAGCACCAAGAAAAATCTTGGCAATTTGGGAAACCTAATGG 827
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTTAATTGATGTTCTTCTAAGACTTATGAATGATGAGGCGCTCAA 887
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrIleGlySer 298
Db 888 TTTCCTATCACCAACGACAAACATCAAGCCATAATTTTGTGACATGTTTGTCTCCGGGACA 947
```


[illegible]

RESULT 12

US-11-116-881A-220
; Sequence 220, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933

Db	828	TTAGGAGTGAAGATTAAATTGATGTTCTTCTTAAGACTTATGAATGATGGAGGCTTCAA	887
Qy	279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	888	TTTCTATCACAACGACAACATCAAAAGCCATAATTTTGTGCTGCCGGACA	947
Qy	299	GluThrThrSerLysAlaThrGluTirPValMetAlaGluLeuMetLysAsnProThrGlu	318
Db	948	GAGACTTCATCGTCAACAATTTGTGTGGCTATGTTAGAAATGGTGAATAATCCAGCCGTA	1007
Qy	319	LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu	338
Db	1008	TTCCGGAAGCTCAAGCAGAAGTAAGAGAAGCAATTTAGAGAAAGAAACTTTTCGATGAA	1067
Qy	339	SerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHis	358
Db	1068	AATGATGTGGAGGAGCTAAACTACTAAAGTTAGTAATAAAGAAACTCTTAAGACTTCAT	1127
Qy	359	ProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgIleAspGlyTyr	377
Db	1128	CCACCGTTCACATTTTGTCTCCCAAGAGAATGTAGGGAAGAGACAAATATAAACGGCTAC	1187
Qy	378	GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn	397
Db	1188	ACTATTCTGTGAAGACCAAGTCAATGTTTAATGTTTGGGCTTTGGGAAGAGATCCAAA	1247
Qy	398	ThrTirSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
Db	1248	TATTGGATGACGCAGAACTTTTATGCCAGAGAGATTTGAGCAGTCTCTAAGANTTTT	1307
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Db	1308	GTTGGTAATAATTTTGAATATCTTCCATTTGGTCGCGAAGAGGATTTGTCTCTGGGATT	1367
Qy	438	ThrSerAlaIleThrAsnLeuLutyrValIleIleAsnLeuLeuTyrHisPheAsnTirp	457
Db	1368	TCGTTTGGCTTAGCTAATAGCTTATTTOCCATTTGGCTCAATTACTATATATCACTTCGATGG	1427
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
Db	1428	AAACTCCTCTGGAATCGAACCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAACT	1487
Qy	478	LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln	492
Db	1488	GCCGCTAGAAAAAGTGACCTTTACTTGGTTGCGACTCCTTATCAA	1532
RESULT 13			
US-11-116-881A-240			
; Sequence 240, Application US/11116881A			
; Publication No. US20060041949A1			
; GENERAL INFORMATION:			
; APPLICANT: Xu, Dongmei			
; APPLICANT: Nielsen, Mark T.			
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof			
; FILE REFERENCE: 07678/141014			
; CURRENT APPLICATION NUMBER: US/11/116,881A			
; CURRENT FILING DATE: 2005-04-27			
; PRIOR APPLICATION NUMBER: 60/665,451			
; PRIOR FILING DATE: 2005-03-24			
; PRIOR APPLICATION NUMBER: 60/665,097			
; PRIOR FILING DATE: 2005-03-24			
; PRIOR APPLICATION NUMBER: 60/646,764			
; PRIOR FILING DATE: 2005-01-25			
; PRIOR APPLICATION NUMBER: 60/607,357			
; PRIOR FILING DATE: 2004-09-03			
; PRIOR APPLICATION NUMBER: 60/566,235			
; PRIOR FILING DATE: 2004-04-29			
; PRIOR APPLICATION NUMBER: 10/934,944			
; PRIOR FILING DATE: 2004-09-03			
; PRIOR APPLICATION NUMBER: 10/943,507			
; PRIOR FILING DATE: 2004-09-17			
; PRIOR APPLICATION NUMBER: 60/503,989			

Db	768	GTTCAGAAATGTCATCAATGAGCACAAGAAAAATCTTGCAAATGGGAAAACTAATGAGCG	827
Qy	259	PheGluAlaAaspAenLeuMetAaspValLeuAenLeuGlnLysAenGlyAenValPro	278
Db	828	TTAGGAGGTGAGNATTAATGTATCTTCTTAAGACTTATGAATGATGGAGCCCTTCAA	887
Qy	279	ValProValThrAenGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	888	TTTCTCTATCACAACGACAACATCAAGCCATAATTTTTCACATGTTGTGTCGGCGACA	947
Qy	299	GluThrThrSerLysAlaThrGluTprValMetAlaGluLeuMetLysAenProThrGlu	318
Db	948	GAGACTTCATCGTCAACAATTTGTGTGGGCTATGTTAGAAATGGTGAATAATCCAGCGGTA	1007
Qy	319	LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAaspGlu	338
Db	1008	TTCCGGAAGCTCAAGCAGAGAAGTAAGAGAAGCATTTAGAGGAAAAAGAACTTTCGATGAA	1067
Qy	339	SerArgPheHisAepLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHis	358
Db	1068	AATGATGTGGAGGAGCTAAATACCTAAAGTTAGTAATAAAGAAGAACTCTTAAGACTTCAT	1127
Qy	359	ProProVal--ValLeuIleProArgGluCysArgGluThrThrArgIleAaspGlyTyr	377
Db	1128	CCACCGGTTCCACTTTTGCTCCCAAGAGAATGTAGGGAAGAGACNAATATAACGGCTAC	1187
Qy	378	GluIleHisProAenThrArgIleValValAsnAlaTrpAlaIleGlyArgAaspProAen	397
Db	1188	ACTATTCTCTGAAGACCAAGTCATGTTAAATGTTTGGGCTTTGGGAAGAGATCCAAAA	1247
Qy	398	ThrTrpSerGluProGlyLysPheAenProGluArgPheLysAaspCysAlaIleAaspTyr	417
Db	1248	TATTGGAAATGACCGAGAAACTTTTATCCAGAGAGATTGAGCAGCTCTCTAAGGATTTT	1307
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyValaGlyLysArgIleCysProGlyIle	437
Db	1308	GTTCGTAATAATTTTGAATATCTTCCATTTTGGTGGCGAAGAGGATTTCTCTGGGATT	1367
Qy	438	ThrSerAlaIleThrAenLeuGluTyrValIleIleAenLeuLeuTyrHisPheAenTrp	457
Db	1368	TCGTTTGGCTTAGCTAATGCTTATTTGCCAATGGCTCAATTACTATATCATCTTCGATTGG	1427
Qy	458	GluLeuAlaAaspGlyIleThrProGlnThrLeuAaspMetThrGluAlaIleGlyGlyAla	477
Db	1428	AAACTCCCTGCTGGAAATCGAACCAAGCGACTTGGACTTGGACTGAGTTGGTTGGAGTA	1487
Qy	478	LeuArgLysLysIleAaspLeuLysLeuIleProIleProTyrGln	492
Db	1488	GCCGCTAGAAAAAGTGACCTTTACTTGGTGGCACTCCTTATCAA	1532

RESULT 14

```

US-10-686-947-209
; Sequence 209, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 209
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-209

```

Alignment Scores:

Pred. No.:

Length: 1673

[illegible]


```

Db      1008 TTCCGCGAAAGCTCAAGCAGAAGTAAGAGAGCATTTAGAGGAAAGAAACTTTTCGATGAA 1067
QY      SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db      1068 AATGATGTGGAGGAGCTAACTACCTAAAGTTAGTAATAAAGAAACTCTTAAGACTTCAT 1127
QY      359 ProProVal---ValLeuLeuProArgGluCyAsArgGluThrThrArgLysAspGlyTyr 377
Db      1128 CCACCGGTTCCACTTTTGCTCCCAAGAGAATGTAGGGAGAGACAATAATAACGGCTAC 1187
QY      378 GluLysHisProAsnThrArgLysLeuValValAsnAlaTIPAlaLysGlyArgAspProAsn 397
Db      1188 ACTATTCTCTGTAAGACCAAGTCACTGGTTAATGTTGGGCTTTGGGAAGAGATCCAAAA 1247
QY      398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaLysAspTyr 417
Db      1248 TATTGGAAATGACCAGAACTTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGAATTT 1307
QY      418 LysGlyThrThrPheGluLeuValProPheGlyValaGlyLysArgLysCysProGlyLys 437
Db      1308 GTTGGTAATAATTTTGAATATCTTCCATTTGGTGGCGGAGGAGGATTTGCTCGGAT 1367
QY      438 ThrSerAlaLysThrAsnLeuGluTyrValLysLeuAsnLeuLeuLysPheAsnTyr 457
Db      1368 TCGTTTGGCTTAGCTAATGCTTATTTGCCATTGGCTCAATTTACTATATCATCATTCGATTG 1427
QY      458 GluLeuAlaAspGlyLysThrProGlnThrLeuAspMetThrGluAlaLysGlyAla 477
Db      1428 AAACCTCCCTGCTGGAATCGAACCAACGACTTGGACTTGGACTTGGTGGAGTAACT 1487
QY      478 LeuArgLysLysLysLeuAspLeuLysLeuLysProLysProLysGln 492
Db      1488 GCGGCTAGAAAAGTGACCTTTACTTGGTTGGGACTCTCTTAACA 1532

RESULT 15
US-10-943-507-206
; Sequence 206, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 Clip - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 206
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-943-507-206

Alignment Scores:
Pred. No.: 21e-119 Length: 1673
Score: 1148.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: 10 Gaps: 5
```

```

US-10-759-813-2 (1-500) x US-10-943-507-206 (1-1673)
QY      12 LeuLysSerPheLeuLeuValLeuLysGluThrLeuArgLeuTyrLysGln 31
Db      54 CTGGTTTCCATTTTCTCTATCTTTCTCTCTTTTAAAGGATGATGAAGAACTCC 113
QY      32 AsnPro-----ProProGlyProTyrLysPheProLysLeuLysGln 46
Db      114 AATAGCCAAAGCAAAAGTTGCCACCAAGTCCATGGAAACTACCAATACCTAGGAAGTATG 173
QY      47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db      174 CTTTCATATGGT-----GGTGGACTACCAACCACTGCTCTTAGAGATTTAGCCAAAAA 227
QY      67 TyrGlyProValMetSerLeuGlnLysGlnValSerAlaValLysSerAla 86
Db      228 TATGGACCACTTATGCACCTTCAATTAGGTGAAGTTTTCGGGTGTGGTTACTTCTCT 287
QY      87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProLysVal 106
Db      288 GATACGGCAAAAGAGTATTAAAACTCATGACATCGCTTTTGGCTAGGCCCTAGCCCT 347
QY      107 LeuAspAlaGlnLysValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db      348 TTGGCCCGGAGATTGCTGTTACAATAGTCTGATCTAGCTTGTGCTTGGCCCTATGCGC 407
QY      127 HisTrpArgGlnMetLysLysLysLysLysLysLysLysLysLysLysValGln 146
Db      408 TATTTGAGACAATCGTAAATAATGTGCTTTGGAAAGTCTCAGTCCCAAGAAATGTCGG 467
QY      147 SerSerArgLeuLysArgGluGluMetGluAspAlaLysLeuThrPheLeuArgSerLys 166
Db      468 ACATTAGTCTCTATTAGCGGAATGAAGTCTCTGCTCATTAATTTTATCCGTCATCT 527
QY      167 AlaGlySerProValAsnLysLysLysLysLysLysLysLysLysLysLysMetIle 186
Db      528 TCTGTGTAACCTTATTAACTTACGGAAGGATCTTTTGTTCACAAAGCTCCATGACATG 587
QY      187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db      588 AGATCAGCGTTTGGGCAAGTGTTCAAAGAGCAAGACAAATTTATACAATTAATAAGAA 647
QY      206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db      648 GTGATACCTTTAGCAGGAGGTTTGCATGTGGCTGACATATTCCCTCCTCAGTGAATTC 707
QY      226 HisTyrLysLysGlyAlaGluSerLysProArgArgLeuHisGlnLysLysLysLys 245
Db      708 CATGTGCTCAGTGAATGAAGGTAAGATTATGAATGCACACCAATAAGGTAGATGCCAT 767
QY      246 LeuGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysPro 258
Db      768 GTTGAGAATGTCATCAATGAGCACAAGAAAAATCTTTGCAATTTGGGAAAAACTAATGAG 827
QY      259 PheGluAlaAspAsnLeuMetAspValLeuLeuLeuGlnLysAsnGlyAsnValPro 278
Db      828 TTAGGAGGTGAAGATTAAATGATGTTCTTCTAAGACTTATGAATCATGATGAGGCCCTCA 887
QY      279 ValProValThrAsnGluSerLysLysLysLysLysLysLysLysLysLysLysLys 298
Db      888 TTTCTTATCACCACGACACATCAAAAGCTATAATTTTTTGACATGTTTCTGCTCGCGGAC 947
QY      299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db      948 GAGACTTTCATCTCAACAAATTTGTGGGCTATGGTGAAGAAATGGTGAAGAAATCCAGCC 1007
QY      319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db      1008 TTTCGCGAAAGCTCAAGCAGAGACTAAGAGAGCATTTAGAGGAAAAAGAAACTTTTCGATGA 1067
QY      339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db      1068 AATGATGTGGAGGAGCTAACTACCTACCTAAGTTAGTAATAAAGAAACTCTTAAGACTTCAT 1127
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 08:37:04 ; Search time 30 Seconds
(without alignments)
2944.349 Million cell updates/sec

Title: US-10-759-813-2
Perfect score: 2574
Sequence: 1 MQQKLSFSPISILFLVLI.....KIDKLIPYQVSLGSNIS 500

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 246837 seqs, 58886990 residues

Total number of hits satisfying chosen parameters: 493674

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DRV=xlh
-Q=/abs/ABSWEB_spool/US10759813/runat_31052006_060316_2859/app_query.fasta_1
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs802h
-USER=US10759813 @CGN_1_18 @runat_31052006_060316_2859 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSBLOK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : Published Applications NA New.*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1099	42.7	1751	6	US-10-953-349-22681 Sequence 22681, A
2	1054.5	41.0	1504	6	US-10-953-349-24167 Sequence 24167, A
3	944	36.7	1632	6	US-10-953-349-32482 Sequence 32482, A
4	943.5	36.7	1766	6	US-10-953-349-32468 Sequence 32468, A
5	911.5	35.4	1711	6	US-10-953-349-33880 Sequence 33880, A
6	840	32.6	1915	6	US-10-953-349-13595 Sequence 13595, A
7	828.5	32.2	1737	6	US-10-953-349-14590 Sequence 14590, A
8	817.5	31.8	1617	6	US-10-953-349-2239 Sequence 2239, Ap
9	816.5	31.7	1586	6	US-10-953-349-4847 Sequence 4847, Ap

10	808	31.4	1720	6	US-10-953-349-32176 Sequence 32176, A
11	801	31.1	1636	6	US-10-953-349-21349 Sequence 21349, A
12	799	31.0	1637	6	US-10-953-349-6466 Sequence 6466, Ap
13	775.5	30.1	1675	6	US-10-953-349-31583 Sequence 31583, A
14	760.5	29.5	1852	6	US-10-953-349-12597 Sequence 12597, A
15	760	29.5	1660	6	US-10-953-349-21323 Sequence 21323, A
16	759.5	29.5	1883	6	US-10-953-349-32448 Sequence 32448, A
17	742.5	28.8	1958	6	US-10-953-349-17424 Sequence 17424, A
18	739.5	28.7	1860	6	US-10-953-349-32215 Sequence 32215, A
19	727.5	28.3	1727	6	US-10-953-349-12509 Sequence 12509, A
20	727	28.2	1772	6	US-10-953-349-34350 Sequence 34350, A
21	718	27.9	1831	6	US-10-953-349-32436 Sequence 32436, A
22	716.5	27.8	1964	6	US-10-953-349-32460 Sequence 32460, A
23	716	27.8	1866	6	US-10-953-349-13642 Sequence 13642, A
24	688.5	26.7	1775	6	US-10-953-349-32421 Sequence 32421, A
25	686.5	26.7	1722	6	US-10-953-349-10369 Sequence 10369, A
26	684.5	26.6	1774	6	US-10-953-349-32754 Sequence 32754, A
27	679.5	26.4	1681	6	US-10-953-349-9959 Sequence 9959, Ap
28	659.5	25.6	1659	6	US-10-953-349-13112 Sequence 13112, A
29	657	25.5	1660	6	US-10-953-349-32409 Sequence 32409, A
30	654.5	25.4	1658	6	US-10-953-349-17280 Sequence 17280, A
31	651	25.3	1639	6	US-10-953-349-32472 Sequence 32472, A
32	644.5	25.0	1725	6	US-10-953-349-32385 Sequence 32385, A
33	644	25.0	2175	6	US-10-953-349-25760 Sequence 25760, A
34	640	24.9	1849	6	US-10-953-349-25544 Sequence 25544, A
35	635.5	24.7	1681	6	US-10-953-349-34955 Sequence 34955, A
36	629	24.4	1828	6	US-10-953-349-34101 Sequence 34101, A
37	598	23.2	1690	6	US-10-953-349-12761 Sequence 12761, A
38	576	22.4	1775	6	US-10-953-349-18767 Sequence 18767, A
39	571	22.2	1756	6	US-10-953-349-13063 Sequence 13063, A
40	570	22.1	1842	6	US-10-953-349-32401 Sequence 32401, A
41	564.5	21.9	1731	6	US-10-953-349-32389 Sequence 32389, A
42	561	21.8	1867	6	US-10-953-349-13128 Sequence 13128, A
43	549.5	21.3	1659	6	US-10-953-349-41434 Sequence 41434, Ap
44	540	21.0	1825	6	US-10-953-349-34905 Sequence 34905, A
45	536.5	20.8	1667	6	US-10-953-349-2736 Sequence 2736, Ap

ALIGNMENTS

RESULT 1
US-10-953-349-22681
; Sequence 22681, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22681
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-22681

Alignment Scores:
Pred. No.: 1.9e-108 Length: 1751
Score: 1099.00 Matches: 224
Percent Similarity: 63.7% Conservatives: 94
Best Local Similarity: 44.9% Mismatches: 163
Query Match: 42.7% Indels: 6
DB: 6 Gaps: 6

US-10-759-813-2 (1-500) x US-10-953-349-22681 (1-1751)

QY 7 SerPheProSerIleLeuIleSerPheLeuValLeuValValMetArg 26
Db 49 TCGTTCCTGTTTATTAGTCTGTTTCTCTCTGTTGTCATTG---CTTCAAGCAT 105
:::
:::
:::

Qy	27	LeuTrpLysLysGlnAsnPro-----ProProGlyProTrpLysPhePro	41
Db	106	TATTACAGCAAAACACCCCTTTCTCACAAACTACCCCGGAGCAAAAGAGCTACCT	165
Qy	42	IleIleGlyAsnLeuProHisLysLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArg	61
Db	166	CTGATAGGGAACCTGCATCAACTAGCAATGGCTGGTTCACTTCCACATCGTACTCTCCGA	225
Qy	62	AlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaVal	81
Db	226	GACCTTGCCCTTAAATACGACCTCTCATGCATCTCCAACCTGGTGAAATTTCTCCGTG	285
Qy	82	ValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnLalaAspAlaPheAla	101
Db	286	GTTGTATCTCCCCCAACATGGCCAGGAGATAATGAAACCCCATGATGATCTGCTTTGTG	345
Qy	102	GlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPhe	121
Db	346	CAGAGGCCCAATTTCTCTGCTCAAAATTTGACCTATGGGCAAAATGATATGTGTTT	405
Qy	122	AlaSerTyrGlyAspHisTyrArgGlnMetLysLysIleTyrIleLeuGluPheLeuSer	141
Db	406	GCTCCATATGGGGATATTGGAGCAAAATGAGAAATAATATGTCTCGGAGCTCTAAGT	465
Qy	142	AlaLysLysValGlnSerSerArgLeuIleArgGlnGluMetGluAspAlaIleThr	161
Db	466	GCCAAAGAGATTCACTTTCTCTCATATTAGAGACGAGACATCCAAATTTATAGAA	525
Qy	162	PheLeuArgSerLysAlaGlySerProValAsnIleThrLysLysIleIleTyrGlyIleIle	181
Db	526	TCAATTCCGAATATCAGAAAGGTTCAACCAATCAATCTTACTAGTAAATTTACTCCTTGTA	585
Qy	182	IleSerIleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeu	200
Db	586	AGTTCTTCTGTTCCAGGTAGCATATTAGCGGGGAGGATTTGAACTGACGATTTGTTTCT	645
Qy	201	SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro	220
Db	646	TGTGTGCTTGAAAAATCATATTAGCGGGGAGGATTTGAACTGACGATTTGTTTCTCT	705
Qy	221	ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln	240
Db	706	TCGATGAAA---CTTCATCTCATAAATGGAAGGAGGCCAAATTTGGAGAAGATGCATGAG	762
Qy	241	GluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys-----	257
Db	763	CAGGTTGCAAGATCGCAGACACACATCTCTAGAGAGCATCAAGAAAGGCCCAAGGGCA	822
Qy	258	-----ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeu	271
Db	823	CTACGAGAGGCAAGGTTGACCTAGAGGAGGAGGATCTGTGTGATGTTCTTTTGAGATC	882
Qy	272	GlnLysAsnGlyAsnValProValThrAsnGluSerIleLysAlaSerValLeu	291
Db	883	CAACAAAGTGCAACACCTCGAGATCAAAATTAGTACCAACAACATCAAGGCTGTATATTG	942
Qy	292	GlnMetPheThrAlaGlySerGluThrThrSerLysLysAlaThrGluTrpValMetAlaGlu	311
Db	943	GACCTATTCTGCTGGACGCTACTTTCAGCATCAACACTTGAATGGGCTATGGCGAA	1002
Qy	312	LeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGly	331
Db	1003	ATGATGAGAAACCCACAGAGTGAGGAGAAAGACACAGGCTGAAGTAAGCAAGCATTTAGG	1062
Qy	332	GluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValVal	351
Db	1063	GAATTGAAATAATTCCTATGAACAGATGTAGGGAAACTTACATCTTTAAAGTTGGTGATC	1122
Qy	352	LysGluThrLeuArgLeuHis---ProProValValLeuIleProArgGluCysArgGlu	370
Db	1123	AAAGAGACATTAAGGTTACACGCTCCTTCTCTTTATTGGTCCCTAGAGAAATGCTCTGAA	1182
Qy	371	ThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrp	390

Db	1183	CTAACCATCATTTGATGGATATGAAATACCTGTCAGACCTAAAGTCATGATGATAAACGTCGTGG	1242
Qy	391	AlaileGlyArgAspProAsenThrTrpSerGluProGlyLysPheAsnProGluuArgpPhe	410
Db	1243	GCAATTTGGAAGAGATCCCAATATTGGCATGATGCTCAGAGAGTTTCTCCACAGAGATTC	1302
Qy	411	LysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGly	430
Db	1303	GATGGTAGTTCTATCGATTTCAAAGGGAATAACTTTTGGTATCTCTCTTTTGGGCGAGGA	1362
Qy	431	LysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsn	450
Db	1363	AGGAGAATGTGCCAGGAATGACATTTGGTTTATAGCAACATTAAGTCTCCACTCGCTCTA	1422
Qy	451	LeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMet	470
Db	1423	TTACTCTATCACTTCAATTTGGGAACCTTCCAAATGAGATGAACCTGAGGATATGGATATG	1482
Qy	471	ThrGluAlaIleGlyGlyAlaLeuArgLysGlyIleAspLeuLysLeuIleProIle	489
Db	1483	TCTGAAAACCTTTGGATTGACAGTTACAAGGAAAAGCGAGTTGTGTTTGATTCCCAT	1539
RESULT 2			
US-10-953-349-24167			
; Sequence 24167, Application US/10953349			
; Publication No. US20060107345A1			
; GENERAL INFORMATION:			
; APPLICANT: ALEXANDROV, Nikolai et al.			
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLY			
; TITLE OF INVENTION: ENCODED THERBY			
; FILE REFERENCE: 2750-1579PUS2			
; CURRENT APPLICATION NUMBER: US/10/953,349			
; CURRENT FILING DATE: 2004-09-30			
; NUMBER OF SEQ ID NOS: 40252			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 24167			
; LENGTH: 1504			
; TYPE: DNA			
; ORGANISM: Glycine max			
US-10-953-349-24167			
Alignment Scores:			
Pred. No.: 8,91e-104 Length: 1504			
Score: 1054.50 Matches: 209			
Percent Similarity: 65.2% Conservative: 79			
Best Local Similarity: 47.3% Mismatches: 145			
Query Match: 41.0% Indels: 9			
DB: 6 Gaps: 4			
US-10-759-813-2 (1-500) x US-10-953-349-24167 (1-1504)			
Qy	59	ArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnVal	78
Db	6	AAATTAAGAGACCTGGCCCATAAATATGGACCCGTGATGATCTTCAACTTGGACAGTT	65
Qy	79	SerAlaValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAsp	98
Db	66	TCACTATTGTCTATTTTCATCCCTGAGTGCTAGGGAAGTGATGAAACCCATGACATT	125
Qy	99	AlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAsp	118
Db	126	AACTTTGCCACAGGCCATAAAGTTCTAGCTATTGAAATAATGTCTTCAACAATTCACAAAGT	185
Qy	119	ValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGlu	138
Db	186	ATAGCTTTTGCTGTGCTATGGAATATTATGGAGGCGAGCTAGAAATAATATGTCATAAATATG	245
Qy	139	PheLeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAsp	158
Db	246	CTTTTAGCCCTAAACCGCTCAACTCATTCAGCCCAATTAGAGAAGATGAGCTCTTCAAT	305
Qy	159	AlaIleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyr	178


```
Db 689 GACATTTTCCCTCCTCGCGTCTCGTGGGCTCCTTAGCAGCGTCCCGCGCAGGATACAG 748
Qy 237 ArgLeuHisGlnGluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsn 256
Db 749 CGCCATAGCAACATGAAGCTGTTTCATGGACACCATCATCAAGAGCACCAGGTGAAC 808
Qy 257 LysProPheGluAlaAsp-----AsnLeuMetAspValLeuLeuAsnLeu 271
Db 809 AGCGGCCCGCAGTCGCGCGGCAGCAAAAGAAAGAGACTTGTCTCGACGTGCTCTGAGACTG 868
Qy 272 GlnLysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeu 291
Db 869 CAAAGGAGGAGACTCGCAGTATCCACTCACCATGACAAATCAAGACCGTCAGTG 928
Qy 292 GlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGlu 311
Db 929 GACATGTTTGGCGCGCGCAGCAGAGCTGCGGCGACGCTGCACTGCGCGATGCGCGAG 988
Qy 312 LeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGly 331
Db 989 CTGATACGGAACCCCGGGTTATGCGGAAGCGCAAGACGAGGTCCGCAGCAACTCGCC 1048
Qy 332 GluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuVal 351
Db 1049 GGGCAGCAGAGTGAAGGAGCGCCACCTGGCAGATCTGCGATCCTTGGGTTTGTATC 1108
Qy 352 LysGluThrLeuArgLysHisProProVal---ValLeuIleProArgGluCysArgGlu 370
Db 1109 AAGGAGAGCGTCGAGGATGACCCCGCGCGCCATTGCTGCTCCCGCTAGGTGCGGGAGC 1168
Qy 371 ThrThrArgIleAspGlyTyroGluLeuHisProAsnThrArgIleValValAsnAlaTrp 390
Db 1169 CCGTGTCAAGTTCTCGGCTCGAGCTGCGGAGGGGGTTCATGGTATCGTGAACCGTGG 1228
Qy 391 AlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPhe 410
Db 1229 GCGATCGGATGACCCCGCGCACTGGGAGGCACCGAGGAGTTCGACCGGAGCGGTTC 1288
Qy 411 LysAspCysAlaIleAspTyroLysGlyThrThrPheGluLeuValProPheGlyAlaGly 430
Db 1289 GAGCAGAACGGGAGGACTTCAAGGGCGGAGTTCGAGTTCTGCTGCGCTTGGCGCGGG 1348
Qy 431 LysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyroValIleIleAsn 450
Db 1349 AGGAGATATGCCCGCGCATGCTTTCGGGCTGCGCGCACGTCGAGCTCGCGCTCGCGCG 1408
Qy 451 LeuLeuTyroHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMet 470
Db 1409 CTGCTGTTCCACTTCGACTGGAGTGCACGCGGGCGGCGCGGCGGAGGATCTGGACATG 1468
Qy 471 ThrGluAlaIleGlyGlyAlaLeuArgLysLysIleAspLeuLysLeuIleProIle 489
Db 1469 ACCGAGGAGTTTGGCGTCACGCGACGGCTCCGGTCTGACCTTGTGCTGCTCGCGCTG 1525
```

RESULT 4

```
US-10-953-349-32468
; Sequence 32468, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32468
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-32468
```

```
Alignment Scores:
Pred. No.: 8.76e-92 Length: 1766
Score: 943.50 Matches: 190
Percent Similarity: 60.9% Conservative: 115
Best Local Similarity: 37.9% Mismatches: 177
Query Match: 36.7% Indels: 19
DB: 6 Gaps: 9
```

US-10-759-813-2 (1-500) x US-10-953-349-32468 (1-1766)

```
Qy 7 SerPheProSerIleLeuIleSerPheLeuLeuValLeuIleLeuValValVal----- 24
Db 83 ACCATTATCAGGTCGTGCTCGCGTAGCTCTGTTTCGCTCATCATGATTCGGGGC 142
Qy 25 MetArgLeuTrpLysLysGlnAsnPro-----ProProGlyProTrpLysPhe 40
Db 143 CGCCGCGCTCGGGCGCGGTGGCGACGGCTGCGTATGGCGCTTGGCGCGTGGCAGCTG 202
Qy 41 ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe 60
Db 203 CCGGTGATCGGCACCTGACACCTCTGCTGCGCGGCGAGCTCCCTCACCGCGCGATG 262
Qy 61 ArgAlaLeuAlaGlnIleTyroGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
Db 263 CGCAGCTGGCGCGCAGCTACGGCGCGCGCTGCTGCTCCAGCTCGGCGAGGTGAAGCG 322
Qy 81 ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe 100
Db 323 CTGTGTGTCTCTCGGAGGGCGCGCGAGGTGATGAAGACCAACACCAACCATGTTC 382
Qy 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyroAsnArgLysAspValLeu 120
Db 383 GCCACGGCGCGCTGAGCACCCACCATGCGCGTGTCTCTACCGCGCGCAGGACATCGTC 442
Qy 121 PheAlaSerTyroGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
Db 443 TTCGGGCTTACGAGAACTACTGGCGCAGCTCCGCAAGATCGCGTGTCCGAGCTCTTC 502
Qy 141 SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAlaIle 160
Db 503 ACCGCGCGCGCTCTCTCTCGCGCCATCGCGAGGAGGAGGTCCCAACCGCGCTC 562
Qy 161 ThrPheLeu-----ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyro 178
Db 563 CGCGTGTGCGCGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTATCC 622
Qy 179 GlyIleIleIleSerIleMetIleArgThrSerValGlyAsn---CysLysGlnLysGlu 197
Db 623 ACCTGTCGTCAGCAGCAGCAGCGCGCGCTCATAGCGCAGCGGTGCGAGGAGCGCAGC 682
Qy 198 ArgLeuLeuSerValAlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAsp 217
Db 683 CGGTTCCTCCGGAGCTCGACCGCATCTGTCAGCTCGCGTGGGGTTTCAACCTGCGCGC 742
Qy 218 AlaPheProThrTrpLysLeuHisTyroIleGlyAlaGluSerLysProArgArg 237
Db 743 ATGTGGCGCTGTCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 802
Qy 238 LeuHisGlnGluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys 257
Db 803 CTGTACACCATGCTCGAGGCGATCGTTGAGGAGCACCTG---GAGAGGACGCGCGCGC 859
Qy 258 ProPheGluAlaAspAsnLeuMetAspValLeuLeuLeuLeuGlnLysAsnGlyAsnVal 277
Db 860 GGAGGCCACGCGGAGGACTTGTCTGACGCTGCTCAAGATCCAGAGGAGGAGGTGCTC 919
Qy 278 ProValProValThrAsnGluSerIleLysAlaSerValLeuLeuGlnMetPheThrAlaGly 297
Db 920 AAGTTTCCATCCACATGGACCGCTCAAGGCCATCATCTTGGACGTATTCTCTCGCGCG 979
Qy 298 SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
Db 980 AGTGAACAACAACCTACGACAATTGAGTGGGCCCATATCATGAGCTGTATCAACAACCCGATG 1039
```

```

QY 318 GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
Db 1040 GCCATGCAGAGCGGACAGATAGGTGGACACAGCTTCCATCTAGTGAACCGGTGCC 1099

QY 338 GluSerArgPheHisLeuLysPhePheLysLeuValLysGluThrLeuArgLeu 357
Db 1100 GAGCAGCGCCCTAAGCGAGCTCCCATACCTCGCTTGTGTCATCCGAGAGAGCTGCGGCTA 1159

QY 358 HisProValValLeuLeuProArgGluCysArgGluThrThrArgLeuAspGlyTyr 377
Db 1160 CACCCGCCCTGCGGTGTGTTCCGCGAGTGCAGGAACCGTGCCAAAGTGCAGGGATAC 1219

QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1220 GAGTGCAGCGGGGACGACGAGGTGTGTCATGCTTGGCGGTAGGCCGCGAGCGC 1279

QY 398 ThrTrpSerGlu---ProGlyLysPheAsnProGluArgPheLysAspCysAla----- 414
Db 1280 TATTGCGCGGACGCGCTGAGGAGTTCCGCGCGGAACGGTTCGAGGAAGCAGCAAG 1339

QY 415 IleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCys 434
Db 1340 GAGACTTTGGGGGTGGTGTGCTTGGCTTTCGCTTGGCGCGCGCGCGAGGATGTGC 1399

QY 435 ProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHis 454
Db 1400 CTGGGATGGCTTGGCTCGCGCGCGTGGAGTCCCGCTGCAAGATGTCTTCCAC 1459

QY 455 PheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIle 474
Db 1460 TTTGACTGGAAG---CCACGAGCGCGGAGCTCGCGCGAGTTCGACATGACAGAGAGCTTC 1516

QY 475 GlyGlyAlaLeuArgLysLysIleAspLeuLysLeu-----IleProIle 489
Db 1517 GGCTCACCGCAGCGCGGACAGACGAGCTCTGTGCGCCCTGTCTTGGCGCTGTCTTCTT 1576

QY 490 Pro 490
Db 1577 CCC 1579

RESULT 5
US-10-953-349-33880
; Sequence 33880, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 33880
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33880

Alignment Scores:
Pred. No.: 2,24e-88 Length: 1711
Score: 911.50 Matches: 191
Percent Similarity: 59.4% Conservative: 105
Best Local Similarity: 38.4% Mismatches: 180
Query Match: 35.4% Indels: 22
DB: 6 Gaps: 8

US-10-759-813-2 (1-500) x US-10-953-349-33880 (1-1711)

QY 9 ProSerIleLeuLeuSerPheLeuLeuValLeuLeuValValValMetArgLeuTrp 28
Db 94 CCCAGCGTCTG-----CTTCTCGTACCTTTCTCGCATCCCTCTACTTCTTC 144

```

```

QY 29 -----LysLysGlnAsnPro-----ProProGlyProTrpLys 39
Db 145 TTGGCCACACAGAGTGGAGCGCACCGGGGGCGCGGGCTCCCGCGGGCGTGGCG 204

QY 40 PheProIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArg 59
Db 205 CTGCGCGTGGTGGGACACCTGCATCAC-----CTAGCCCGGGGCGCTCCCGCACCGGCTC 258

QY 60 PheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSer 79
Db 259 ATGCGCGACCTGGCGAGGCGCCACGCGCGCTCATGTCTCGGTTTCGCGAGGTCCCC 318

QY 80 AlaValValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAla 99
Db 319 GTGGTGTGGCTCTCTCGCGCGCGCGCGGAGGTGATGCGGACCCACGAGCGGCG 378

QY 100 PheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspVal 119
Db 379 TTGCGGTGCGGCCCATCGTCCCGTCTCGCGCTCTGGTTCCAGGCGCGCGAGGCGATC 438

QY 120 LeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPhe 139
Db 439 CTCTTTCGCGCCCTACGCGGACGACTGGCGCCACCTCCGAGGGTCTGCACCCAGGAGCTC 498

QY 140 LeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAla 159
Db 499 CTCACCGCGCGCGCGTCCAACTCTTCCCGCGCGCGGAGGAGGAGCTCCCGCGCGCTG 558

QY 160 IleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGly 179
Db 559 CTGCGCTCGTCCGCTCGACGCTCGGG---CCCGTCAACTGACCGAGAGAATATCGACC 615

QY 180 IleIleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArg--- 198
Db 616 TACATCGCGGACTCCACGCTGGCGCCATCATCGGCGCGCGAGGCTCAAGGACCGCGAC 675

QY 199 ---LeuLeuSerValAlaAspAlaValAlaAsnGluAlaAlaThrSerPheGlyThrAlaAsp 217
Db 676 GGTACCTGCGGATGCTCAAGGGCTCTTCGGCATCATCCCGGGATGAGCCTGCGCGAC 735

QY 218 AlaPheProThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArg 237
Db 736 CTCTTCCGCTCTCGCGCTCGCCATGCTCTCTCAGCGCGCTCTCTCCCGGATCCAGGCC 795

QY 238 LeuHisGlnGluIleAspAsp-IleLeuGluGluIleLeuAsnGluHisLys----- 254
Db 796 TACCGTCCGAGCATGGCGAGCATCATGGACGGCATCTCCAGGAGCACCGGACAGAGC 855

QY 255 -AlaAsnLysProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAs 274
Db 856 AGCGCTGGCGAGCGGAGAGAGGACTTCGTTGACGTGCTCTCTCAGACTTCAGAGGA 915

QY 274 nGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPh 294
Db 916 AGTGGACTCCCGATTTCCCGCTCACAACCTGAGAACATCAAAACAGTCTATGCTGGACATCT 975

QY 294 eThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLys 314
Db 976 TGGCGGAGCACGAGAGCTCGACGACGCGCTGGATTGGCGGATGGCGAGCTCTCTGCG 1035

QY 314 sAsnProThrGluLeuLeuGlyAlaGlnGluValArgGlnValPheGlyGluMetG 334
Db 1036 GAAACCCAGGTGATGGAGAGCGCGAGGTGAGGTCCCGCAAGCGCTTTCGCGCCACCG 1095

QY 334 yLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValLysGluThr 354
Db 1096 CGCGGTGACCGAGGACAGACTAGCGCGCTGCGCTACCTGCTGCTCATCAAGAGGATC 1155

QY 354 rLeuArgLeuHisProPro---ValValLeuIleProArgGluCysArgGluThrThrAr 373
Db 1156 GCTGCGGTGACCCCGCGGCGACGATGCTGTTCCAGCCAGTGCAGAGCGCGGTGCCA 1215

```

```
QY 373 gileaspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleG1 393
DB 1216 GGTGTCGGGTACGACGTCCCGCGCGGCATCACGGTGATCGTGAACGGGTGGCGCATCGG 1275
QY 393 yArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCy 413
DB 1276 CNGGACCCCGCCCACTGGGACGACCGACCAAGTCTTTCGCGAAGAGTTTCGACAGAG 1335
QY 413 aAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgI1 433
DB 1336 TACCAGGACITTCAGGGGGCAGACTTCGAGTTTCATACCGTTTCGTCGGCGCGCGCAT 1395
QY 433 eCyAsProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyr 453
DB 1396 CTGCCCCGCGCATGACGTTTCGGCTCGCTCATATCGAGATCGCGCTCGCGCGCTTCTGTT 1455
QY 453 rHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAl 473
DB 1456 CCACITTCGACTGAGTTTCCCGGAGGGCTACCGCGGAGGAGCTGGACATGACCGCAGGC 1515
QY 473 alleGlyGlyAlaLeuArgLysIleAspLeuLysLeuIleProIlePro 490
DB 1516 ATTGGGATTGTACGCCACGCGTTCCGACCTTTTGTGGTGGCGCAACCCCT 1567

RESULT 6
US-10-953-349-13595
; Sequence 13595, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13595
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-13595

Alignment Scores:
Pred. No.: 1.21e-80 Length: 1915
Score: 840.00 Matches: 181
Percent Similarity: 56.7% Conservative: 101
Best Local Similarity: 36.4% Mismatches: 195
Query Match: 32.6% Indels: 20
DB: 6 Gaps: 9

US-10-759-813-2 (1-500) x US-10-953-349-13595 (1-1915)
QY 11 IleLeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLys 30
DB 64 CTAACAGTTTCTTCGCACTTCCTTATATCTTCTACGCTAATAAGAAGA---AACAAA 120
QY 31 GlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeu 50
DB 121 TACAATCTGCCACCATCCCCACCAAGATACCCATAATCGGCAATCTTCACCAGCTA-- 177
QY 51 LeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGlyProVal 70
DB 178 -----GGCACATGCCACACCGCTCTTTCATGCACTCTCACAAAATATGCGCCTCTC 231
QY 71 MetSerLeuGlnIleGlnValSerAlaValIleSerSerAlaGlnAlaLys 90
DB 232 ATGATGTTGCAATGGTCAAAATCCAAACCTTAGTGGTCTCATAGCTGACGTGCGCCAGA 291
QY 91 GluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGln 110
DB 292 GAAATAATCAAAACGCGATGATGTTGTTTCTCCACCGCGGCAACCTACAGCTGCTGTA 351
```

```
QY 111 IleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrpArgGln 130
DB 352 ATCTTTGTTATGATGCAAAAGATGTGGCTTTCTGTGTACTACCGCAAGAGTGGAGACAA 411
QY 131 MetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSerArgLeu 150
DB 412 AAGATAAAGACATGTAGGTTGAGCTTATGATCTGAAGAAGGTGCGGTGTTTTCATCC 471
QY 151 IleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGlySerPro 170
DB 472 ATTAGACAAGAAGTTGTTACAGAGTTGGTGAAGCTATAGGTGAAGCGCTGTGTAGTGAA 531
QY 171 -----ValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIleArg 187
DB 532 AGACCATGTGTGAATCTGACTGAGATGCTGATGGCAGCATCGAACGACATGTGTGTAGA 591
QY 188 ThrSerValGly---AsnCyLysGln-----LysGluArgLeuLeu 200
DB 592 TGTGTTCTTGACCGGAAGTGTGATGTCATGTCATGTTGGTGTGTGGCAGTAGCAGCTTGA 651
QY 201 SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro 220
DB 652 CGGTTGGGACAAAGATTATCAGACTATTATCGGCTTTTCAGCGTGGGTGATTTCTTCCT 711
QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
DB 712 TCGTTGGGTTGGTGTGACTATCTGCTGCTGCTTAATTCAGAGATGAAACACCGCTTCTC 771
QY 241 GluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLysPheProGlu 260
DB 772 GCAGTAGATGTTCTTCTTGATGAGTAATTCGACAGACACGAGCAGCAACAAG---AAG 828
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
DB 829 AATGATGACTTCTTGGGGATACTTCTTCAACTTCAAGAATGTGGGAGGCTTGGACTTTC 888
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
DB 889 CTCGACCGAGATAACCTCAAAGCAATCCTAGTGGACATGATAATAAGTGGGAGTGCACAT 948
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
DB 949 ACTTCAACAACCTCTAGATGACCTTTTCGGGAGTTCTTAGAAATCCAAATACCATGAAG 1008
QY 321 LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLys-----ValAspGlu 338
DB 1009 AAAGCTCAAGAAGAGGTAAAGAAGTGTGGGAAATCAATTCCAAAGCAGTACTGATGAA 1068
QY 339 SerArgPheHisAspLeuLysPhePheLysValValLysGluThrLeuArgLeuHis 358
DB 1069 AATTGTGGAATCAATGAACCTCTTGAATGTGTAGTCAAGAANAACCTTGGATTACAT 1128
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
DB 1129 CCACCCCTCTCTTTCGATTGCTCGAGACATCATCAAGTGTAAATAAGAGGGTAC 1188
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
DB 1189 GATATTCGCGCAAAACAATGGTATTTATCAATGTCATGGCGATCCAGAGGGATCCTGAA 1248
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
DB 1249 TTATGGGATGATCTCGAAGATTTATTCGGAAGATTTGAAACTAGCCAACTGATCTT 1308
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
DB 1309 AATGACAAGATTTTCAATTAATTCGTTCCGTTATGGGAGAGGGGATGCCCTGCAATG 1368
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyrHisPheAsnTrp 457
DB 1369 TCATTGACCTGCTTCAACTGAGTATGTTCTGTAATCTTTTGTATGGTTCATTTGG 1428
QY 458 GluLeuAlaAspGly-----IleThrProGlnThrLeuAspMetThrGluAlaIleGly 475
```



```
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2239
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2239

Alignment Scores:
Pred. No.:      2,43e-78      Length:      1617
Score:          817.50      Matches:      177
Percent Similarity: 55.9%      Conservatives: 111
Best Local Similarity: 34.4%      Mismatches: 192
Query Match:      31.8%      Indels:      35
DB:               6         Gaps:       10

US-10-759-813-2 (1-500) x US-10-953-349-2239 (1-1617)

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
Db 10 TTAATAACAAATAACATACAAAATGATAGCATTTGGATTCTATCATCTCTCTCTTCATC 69
QY 21 LeuValValMetArgLeuTrpLysGlnAsn----- 32
Db 70 ATCTTTCTCTCTCTGCGCCCTTCAACGCAAAATCATGGAAACATCGAGCAATACCA 129
QY 33 ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuLeuThr 52
Db 130 TCTCTCTCTGTT-----TTTCCATCATCGGAATTTATCATCAGCTC----- 171
QY 53 SerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSer 72
Db 172 GGAGAAATTACAAATCAGCTCTATGGAAGCTCTCAAGAAGATATGGCCCTGTGTGCTT 231
QY 73 LeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAlaAlaLysGluVal 92
Db 232 TTGAAGCTTGGAAAAGTCCCCACATCATATCTTCTCATCAGAAACAGCAAAACAAGCT 291
QY 93 MetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleVal 112
Db 292 CTAAGAGACTATGATCTCCATTGTTGTAGCGCTCTCTTAGCAGGAGGAGAGAGCTC 351
QY 113 PheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLys 132
Db 352 TCTTACAACAATCTAGATATGCTTCATCTCCTTATAATGATATATGGAAGAACTAAGG 411
QY 133 LysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSerArgLeuIleArg 152
Db 412 AAGCTCTGAGTCAAGAACTCTTTAGTGTCTATAAATAATCAATCATCAACCTATTAG 471
QY 153 GluGluGluMetGluAspAlaIleThrPheLeu-----ArgSerLysAlaGlySerPro 170
Db 472 GACGAGGAGGTCAAGAAAGTTATCGATTCAATCGCGGAATCATCTTCTTAAAGAATCCG 531
QY 171 ValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIleArgThrSerVal 190
Db 532 GTTAACCTTGAGCAAGACGTTCTTCTGTTAACTACAAGTGTAGTATGCAAGGCAGCATTT 591
QY 191 -----GlyAsnCysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 592 GGTGTGAGTTTGGAGGAAGTGTGCTCAACAGTGTATGATTTCAATAAGTTAGTCCGAGAT 651
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro-----ThrTrpLys 223
Db 652 ACTTTCGAGATGTTGGGAAGCTTCTCGCTCAGATTTTATTCCTGTATGTCGATGG--- 708
QY 224 LeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAsp 243
Db 709 ATCATCGTAAGTTCAATGGTTTACAAGGGTGGAGAAAGAAAGCTTTTCGAGATCTGAT 768
QY 244 AspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLysProPheGluAlaAspAsn 263
```

```
Db 769 CGTTTCTATGAACAAATTTTGTGATCTGCATGAAGAGAGAAAGTAGTAGAAGTCAAGAT 828
QY 264 LeuMetAspValLeuLeuAsnLeuGlnLys-----AsnGlyAsnVal 277
Db 829 TTAGTGGATGTCTCTTGGAGTTGGAGAAAGAAATTTGTTGTGGAAATGGCAAG--- 885
QY 278 ProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db 886 -----CTCACAAAGAAATCATATCAAGCAATTTTGTAGAACATCTCTTTTAGGAGA 936
QY 298 SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
Db 937 ATCGATACTTCTGCAATAACAATGACATGGCAATGGCAGAACCTTGTCTAAAAACCTAGA 996
QY 318 GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
Db 997 GTGATGAAGAAAGTTCAAGCAGAAATCAGAAACCAAAATCAAGAAACAAAGAAATCAGC 1056
QY 338 GluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeu 357
Db 1057 TTTGATGACATGATAAGCTCGAGTACTTGAATGGTGAATGGTGAATGGAGGTTA 1116
QY 358 HisProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGly 376
Db 1117 CATCTCCAAACACCTCTCTCTCCCAAGAGATGTAATCACTGAATTTGAGATCAATGGC 1176
QY 377 TyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspPro 396
Db 1177 TACAGATTCTGCGCAAAACACGCGTTCATGTGAATGTTGGGCTATCGGCGGTATCCT 1236
QY 397 AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAsp 416
Db 1237 GATACATGGAAAGATCCAGAAATGTTCTCCCGAAAGGTTTAAATGATAGTAACATTGAT 1296
QY 417 TyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGly 436
Db 1297 GCAAAAGGACAGAACTTTGAGTTGTTGTCGTTGGAGTGTGTAGGAGAATTTGCTCTGA 1356
QY 437 IleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsn 456
Db 1357 CTGTACATGGGAACAACAATGGTGGAGTTTGGCCTAGCTAATATGTTGTATCATTTTGT 1416
QY 457 TrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGly 476
Db 1417 TGGAAAGTTACAGAAAGGATGTTAGTCGAAGATATCGATATGGAAGAAAGCTCTGGACTT 1476
QY 477 AlaLeuArgLysLysIleAspLysLeuIleProIleProTyr 491
Db 1477 ACTGTGAGCAAAAAAAGTGAGCTTGTACTTGTTCAGTGAAGTAT 1521

RESULT 9
US-10-953-349-4847
; Sequence 4847, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4847
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-4847

Alignment Scores:
Pred. No.:      3.02e-78      Length:      1586
Score:          816.50      Matches:      193
Percent Similarity: 57.4%
```



```
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaLeuIleTyrGly 68
Db 205 ATC-----GGGCCACCGCACCGAGCCTGGCTGGCTGGCCCGCGGCGATGG 255
Qy 69 ProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGluAla 88
Db 256 CCGTGATGATGCTCGCGCTGGCATGGTGGCGACGGTGGTGGTACATCCGACAGAGCG 315
Qy 89 AlalysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAsp 108
Db 316 GCGCGGGAAGCCCTCAAGACCCAGATGATGATGATGATGATGATGATGATGATGATGAT 375
Qy 109 AlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrp 128
Db 376 CCAGGGCTGCTGCTTACCGGTACAAAGACGCTGGCTTCTCGCCGTGGAGGACACTACGTC 435
Qy 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysValGlnSerSer 148
Db 436 CGCGAGATGCCAAGTTGTTTCATCATCGAGTGTCTCAGCAGCGCGCGCTGCAGTCTGCC 495
Qy 149 ArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
Db 496 TACTACGGGAGGATGCACAGATGCACAGCTGATTGAGACCTCACCGTGGTGGGCGG 555
Qy 169 SerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIleArgThr 188
Db 556 AACCGGTACCCCTGGAGGCCACATCTTCGCCACCATGGACCGGATTGTGGCTGTTC 615
Qy 189 SerValGlyAsnCys-----LysGlnLysGluArgLeuLeuSerValAlaAsp 204
Db 616 GCGTTCGGCGAGAGCTACCGGGGAGCAGTTCAAGGGCGAGTGTGTGCCCTTACTCAAC 675
Qy 205 AlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro-----ThrTrp 222
Db 676 GCACCATGGACATGTTGGTGTAGCTTCTGCCCAGACTTCTTCCCAACGCCGTTGGC 735
Qy 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIle 242
Db 736 CGCCTCATCGACCGCATCACGGCGTCAAGGCCACCGCCAGAGGGTCTTTCGTGAGTGC 795
Qy 243 AspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLysProPhe-----Glu 260
Db 796 GACCGCTTCTTTGAGCATATCATCAGCAGTGCATGGGAGAAAGCCACCGCGGCGGC 855
Qy 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsn-----GlyAsnValPro 278
Db 856 GGATCGGACCTGGTCAGAGCTGTGGACATCATGAAGAGCCCGCGCTCCGCTGCA 915
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 916 GGAACCTTCACAGAGATCATGTCAAGGCCATTCTCATGAACACATTCATTGGTAGCAT 975
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 976 GACACCAACACAGTAGTACCATCTGGCAATGGCAGAGCTGATCCGGAACCAAGGTT 1035
Qy 319 LeuArgLysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1036 CTGCAAGGGGCAACTTGATGATTAGACTGCCGTGGAGGGGAGGAGGAGTGCACCA 1095
Qy 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1096 GCTGACATGCCCCAAAATGAGTACCTGAGATGGTGTGTCCGAGACCTACCGGTGCAT 1155
Qy 359 ProPro-----ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrp 377
Db 1156 CCCCCAGCAACACTACTCTGATCCGAGAGAGACGCTCGCGCCGATCCAGGTGGCTGGCTAC 1215
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTyrAlaIleGlyValArgAspProAsn 397
Db 1216 GACATCCCGACCGAACCCAGGTGCTATTGTCAATGTGTGGGCCATCAGCAGGAGACCTTCC 1275
```

```
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1276 GCGTGGGAAGGACCCCGAGGAGTTTAACCCGAGCGGTTCCAGGACACTGACGTGGACTTC 1335
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1336 AACGGCAGCCACTTCGAGTTTCATCCCTTCGCGCGGGCCCGCGGATCTGCCCGGGCTG 1395
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1396 GCCATGGGGTGGCAACGCGAGTACATCTTGCACACTTGTCTACTGTCTCACTGG 1455
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1456 GCGTGGCCCAACGCGGTGAGCGCGGAGTGTGAACATGAGGAGGAAAGGGGTGCTCACT 1515
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyr 491
Db 1516 TACCGGAAGAAGACCTCACTCATCTGCTGTCACACACCTTAC 1557

RESULT 11
US-10-953-349-21349
; Sequence 21349, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21349
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-21349

Alignment Scores:
Pred. No.: 1,45e-76 Length: 1636
Score: 801.00 Matches: 171
Percent Similarity: 56.0% Conservative: 114
Best Local Similarity: 33.6% Mismatches: 184
Query Match: 31.1% Indels: 40
DB: 6 Gaps: 11

US-10-759-813-2 (1-500) x US-10-953-349-21349 (1-1636)

Qy 10 SerIleLeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLys 29
Db 118 ACAATCTTAATCTCTCGTCCCACTAACTACTCTCTGCTGCTTATCTCGTACCCGT 177
Qy 30 LysGlnAsnPro---ProGlyProTrpLysPheProIleIleGlyAsnLeuProHis 48
Db 178 CCAAGACGCGCCTATCCACAGGCGCTTCCCAATCATAGGAAC-----228
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGly 68
Db 229 ATGTTATGATGGACAGCTAACCCACCGCGGTCTCGCCAACTTGGCCAAACACTACGCG 288
Qy 69 ProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAla 88
Db 289 GGAATCTTCCACCTTCGATGGGTTCCTCCACATGTCGCATCTCCGACCCCGCATGCC 348
Qy 89 AlalysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAsp 108
Db 349 GCGGACAGGTTCTCAAAGTCCAGACAAACATCTTTTCCAAACCGCCCGCCACCATCGCC 408
Qy 109 AlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrp 128
Db 409 ATCAGCTACCTCACTACGACCGCGCCGACATGGCTTCGCGCCACTACGCGCCCTTCTGG 468
```

```
QY 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
Db 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
QY 469 CGCCAGATGAGGAACTCTGCGTCATGAAGCTCTTCAGCGCAAGCGCGCGAGTCTCTGG 528
Db 469 CGCCAGATGAGGAACTCTGCGTCATGAAGCTCTTCAGCGCAAGCGCGCGAGTCTCTGG 528
QY 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
Db 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
QY 529 CAGTCGCTTGC---GACGAGGTGCGACTCGCGCGTCCGCCGCGCGCGCGCGCGCG 585
Db 529 CAGTCGCTTGC---GACGAGGTGCGACTCGCGCGTCCGCCGCGCGCGCGCGCGCG 585
QY 169 SerProValAsnIleThrLysIleIleTrpGlyIleIleIleSerIleMetIleArgThr 188
Db 169 SerProValAsnIleThrLysIleIleTrpGlyIleIleIleSerIleMetIleArgThr 188
QY 586 AAACCGCTCAACATCGGAGATGATGTTAACTTCAACAGAACATCATCTACCGCGCC 645
Db 586 AAACCGCTCAACATCGGAGATGATGTTAACTTCAACAGAACATCATCTACCGCGCC 645
QY 189 SerValGlyAsnCysLysGln-----LysGluArgLeuLeuSerValAlaAspAlaVal 206
Db 189 SerValGlyAsnCysLysGln-----LysGluArgLeuLeuSerValAlaAspAlaVal 206
QY 646 CGGTTCCGGTTCGAGTTCGAGGAGGTCAAGAGCATTTTCAATAAATATATTCAGGAGTTC 705
Db 646 CGGTTCCGGTTCGAGTTCGAGGAGGTCAAGAGCATTTTCAATAAATATATTCAGGAGTTC 705
QY 207 AsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHis 226
Db 207 AsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHis 226
QY 706 TCTAAGCTCTTTGGCGCGTTAATATTGCGGATTTTATACCC----- 747
Db 706 TCTAAGCTCTTTGGCGCGTTAATATTGCGGATTTTATACCC----- 747
QY 227 TyrIle-----IleGlyAlaGluSerLysProArgArgLeuHisGlnGlu 241
Db 227 TyrIle-----IleGlyAlaGluSerLysProArgArgLeuHisGlnGlu 241
QY 748 TACCTCGGCGGTAGATCCACAGGTTTGAATCTCAGACTCGCTAGGCGACGTGCGCG 807
Db 748 TACCTCGGCGGTAGATCCACAGGTTTGAATCTCAGACTCGCTAGGCGACGTGCGCG 807
QY 242 IleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys----- 257
Db 242 IleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys----- 257
QY 808 CTCGATAGTTCATTGATAAGATCATCGACGACGATCGCAGAGAAGAATAATATATCAG 867
Db 808 CTCGATAGTTCATTGATAAGATCATCGACGACGATCGCAGAGAAGAATAATATATCAG 867
QY 258 -----ProPheGluAlaAspAsnLeuMetAspValLeu----- 268
Db 258 -----ProPheGluAlaAspAsnLeuMetAspValLeu----- 268
QY 868 AGCAGTGAAATTTGGTATGTTGAAACGGACATGTTGGATGAGTGTGCGGTTCTACGCG 927
Db 868 AGCAGTGAAATTTGGTATGTTGAAACGGACATGTTGGATGAGTGTGCGGTTCTACGCG 927
QY 269 -----LeuAsnLeuGlnLysAsnGlyAsnVal-----ProValProValThr 282
Db 269 -----LeuAsnLeuGlnLysAsnGlyAsnVal-----ProValProValThr 282
QY 928 GAGGAGGCGAAGTTGAATCGGACGATAATTTGCAGAACTCTATCAGACTCACT 987
Db 928 GAGGAGGCGAAGTTGAATCGGACGATAATTTGCAGAACTCTATCAGACTCACT 987
QY 283 AsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSer 302
Db 283 AsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSer 302
QY 988 AAGGATAAATCAAGCTATCATTTATGAGCGTGTGTTTCGGAGGCTCGGAACGGTGGCG 1047
Db 988 AAGGATAAATCAAGCTATCATTTATGAGCGTGTGTTTCGGAGGCTCGGAACGGTGGCG 1047
QY 303 LysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAla 322
Db 303 LysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAla 322
QY 1048 TCAGCAATCGATGGGTCTATGTCGAGCTGATGAGAAGCCAGAGAACCAAAAGCGGTC 1107
Db 1048 TCAGCAATCGATGGGTCTATGTCGAGCTGATGAGAAGCCAGAGAACCAAAAGCGGTC 1107
QY 323 GlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHis 342
Db 323 GlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHis 342
QY 1108 CAACAGAGCTCGCGATGTAGTGGCCCTGGACCGTGGCGGTGGAAGATCCGATTCGAG 1167
Db 1108 CAACAGAGCTCGCGATGTAGTGGCCCTGGACCGTGGCGGTGGAAGATCCGATTCGAG 1167
QY 343 AspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProValVal 362
Db 343 AspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProValVal 362
QY 1168 AAACCTCACTTATCTCAATGTGCCCTCAAAGAGACCTCCGCTCCACCTCCGATACCG 1227
Db 1168 AAACCTCACTTATCTCAATGTGCCCTCAAAGAGACCTCCGCTCCACCTCCGATACCG 1227
QY 363 LeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsn 382
Db 363 LeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsn 382
QY 1228 CTCCTCTCCAGCAGCGCGGAGGACCGCCACCGTGGCGGTACTTCGTCGCCGGAAG 1287
Db 1228 CTCCTCTCCAGCAGCGCGGAGGACCGCCACCGTGGCGGTACTTCGTCGCCGGAAG 1287
QY 383 ThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluPro 402
Db 383 ThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluPro 402
QY 1288 GCGCGTGTATGATCAACCGTGGGCCATTTGGAGGAGCAAGATTCCTGGGAGGAACCC 1347
Db 1288 GCGCGTGTATGATCAACCGTGGGCCATTTGGAGGAGCAAGATTCCTGGGAGGAACCC 1347
QY 403 GlyLysPheAsnProGluArgPhe---LysAspCysAlaIleAspTyrLysGlyThrThr 421
Db 403 GlyLysPheAsnProGluArgPhe---LysAspCysAlaIleAspTyrLysGlyThrThr 421
QY 1348 GAAACTTTTAAAGCCCGCGTTCCTTAAACCGGCGGTGCCCGATTTCAAGGGAGGACAC 1407
Db 1348 GAAACTTTTAAAGCCCGCGTTCCTTAAACCGGCGGTGCCCGATTTCAAGGGAGGACAC 1407
QY 422 PheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIle 441
Db 422 PheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIle 441
QY 1408 TTCAGATTCATTCATTCGCGTGGGTGCAAGATCCTGCCCGGAATGTTGCGGCTC 1467
Db 1408 TTCAGATTCATTCATTCGCGTGGGTGCAAGATCCTGCCCGGAATGTTGCGGCTC 1467
QY 442 ThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAsp 461
Db 442 ThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAsp 461
QY 1468 TAGCGCTTGAATGCGGTGGCTCATCTTCTACCTGCTTACGCGGGAATTCGAGAT 1527
Db 1468 TAGCGCTTGAATGCGGTGGCTCATCTTCTACCTGCTTACGCGGGAATTCGAGAT 1527
QY 462 GlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLysLys 481
```

```
Db 1528 GGAATGAAACCAAGTGAGATGACATGGGTGACGTGTTCCGA-----CTCACCGCTCCA 1581
QY 482 IleAspLeuLysLeuIleProIlePro 490
Db 1582 AGGTCCAGCGACTCATTTGCTGTACCA 1608

RESULT 12
US-10-953-349-6466
; Sequence 6466, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6466
; LENGTH: 1637
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-6466

Alignment Scores:
Pred. No.: 2,378-76 Length: 1637
Score: 799.00 Matches: 178
Percent Similarity: 56.3% Conservative: 106
Best Local Similarity: 35.3% Mismatches: 192
Query Match: 31.0% Indels: 28
DB: Gaps: 11

US-10-759-813-2 (1-500) x US-10-953-349-6466 (1-1637)
```

```
QY 13 IleSerPheLeuValLeuIleLeuValValMetArgLeuTrp----- 28
Db 26 ATCTCTTCTCTGTTGTTTCTCATCACCCTTCGTTTCGTTTAATCTTTTGGCAAGAAA 85
QY 29 -----LysLysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsn 45
Db 86 ATCAACAGCATCAAAATGGAATCTTCTCCCAAGCCCTCCCAAGTTTCCGTCATCGCGAAC 145
QY 46 LeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
Db 146 TTACATCAGATT-----GGAAATTCCTCAGAGTCACCTCAACATCTCGCCGAA 196
QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
Db 197 AGATACGAGACCTGTGATGCTTCTTACCTTGGGTTTGTCCCTATAACTGTGGTCTCATCG 256
QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
Db 257 AGAGAAGCGCGCTGAAGAAGTGTAGAACTCATGACCTAGACTGTGTGACGAGCGCTAAG 316
QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
Db 317 CTGTGGGACAAAGGTACTCTCGCGGATTTAAAGATATCGGTTTACGCCATACGGT 376
QY 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
Db 377 AACGAGTGAAGCGCGCGTAAAGTTTGCCTCGGTGAGCTTTCTGTTTGAAGAGTT 436
QY 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSer 165
Db 437 CAGTCCTTTAGGCATATCCGAGGAGAAATTAATCTTCTGCTCAAGCAACTGTCCGAA 496
QY 166 LysAla-----GlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 497 TCTGCGGTGATGCGCTCTCGGTCGATTTGAGCAAAATCCCTTTCTGCGTAAACCGCTAG 556
QY 184 IleMetIleArgThrSerValGly-----AsnCysLysGlnLysGlu 197
```

```
Db 557 ATCCCTTTTAGAGTTGCTTAGGACAGAAATTTTCACGAGACGATTTTATCGATAAAGAA 616
Qy 198 ArgLeuSerValAlaAspAlaValAaenGluAlaThrSerPheGlyThrAlaAsp 217
Db 617 AAGATCGAAGAGCTGCTTGAAGCTGAGACTGCCCTAGCAAGTTTACACTTGTCTGAT 676
Qy 218 AlaPheProThr-----TrpLysLeuLeuHisTyrIleIleGlyAlaGluSer 233
Db 677 TTTCTTCCCTGTGCGCGAGCTTGGATGG--CTCGTTGATTGGTTTTCGGGACAAACACAAG 733
Qy 234 LysProArgArgLeuHisGlnGluLeuAspIleLeuGluGluLeuLeuAsnGluHis 253
Db 734 AGACTCAACGATGTTTTTACAAAGTCGATGCTCTGTTTCAACATGTCATAGATGATCAT 793
Qy 254 ---LysAlaAsnLysProPheGluAlaAspAsnLeuMetAspValLeuLeuAsn---Leu 271
Db 794 TTAAATCTGGNAGATCAAAAGAGCAGCAGACATCATCGATTCAATGTTGGATGTGATT 853
Qy 272 GlnLysAsnGly---AsnValProValProValThrAsnGluSerIleLysAlaSerVal 290
Db 854 CATAAACAAGGAGAGACAGTTCCTTAGAGCTCACATAGATCATATCAAGGGGTTTCTC 913
Qy 291 LeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAla 310
Db 914 GCGAATATATTTCTTGCAGGGGATGACACACAGGGGCCATCACCATGATATGGCAGTGACG 973
Qy 311 GluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluGluValArgGlnValPhe 330
Db 974 GAGCTCGTTTAAACCCGAACTGATAAAGAAAGTTTCAAGGCGATATCGGAGAACACTT 1033
Qy 331 Gly---GluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeu 349
Db 1034 GGCAGCAATAAGGAGAGAATACCCGAGGAAGATATCGAGAAAGTTTCTTACTTGAAGATG 1093
Qy 350 ValValLysGluThrLeuArgLeuHisProProValValLeuLe---ProArgGluCys 368
Db 1094 GTAACTAAGAAACAACTTACAGTTTACACCCAGCAGCTCCTCTTATATCTTCCAAAGGGAACA 1153
Qy 369 ArgGluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAaen 388
Db 1154 ATGGCTCACATCAAGTTTCAAGGGTATGATATTTCTCCCAAGAGGAGATCTTGTGCAAT 1213
Qy 389 AlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGlu 408
Db 1214 GTTTCGGCAATAGGAAGAGATCCCAACTCTGACAAACCCGAAAGAGTTTGACCTGAG 1273
Qy 409 ArgPheLysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGly 428
Db 1274 AGGTTTATGGATAGCTTTGTTGATTATAGGGACAAACATTACGAGCTCTTACCAATTGGG 1333
Qy 429 AlaGlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIle 448
Db 1334 TCCGGTCGAAGATATGTTCCCGGATGCCAATGGGATTTGCTCCGTCGAATTTGGACTC 1393
Qy 449 IleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeu 468
Db 1394 TTGAACCTTACTTCTTCTGATTGGAAGTTGCCTGATGGATGACACATAAAGATATC 1453
Qy 469 AspMetThrGluAlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeuIlePro 488
Db 1454 GATACTGAAGAGCTGGTACTCTTACAATAGTCAAGAAAGTACCTCTCAAGCTCGTTCCA 1513
Qy 489 IleProTyrGln 492
Db 1514 GTTCGAGTTTCAG 1525
```

RESULT 13

```
US-10-953-349-31583
; Sequence 31583, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
```

```
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THEREY
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31583
; LENGTH: 1675
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-31583
```

```
Alignment Scores:
Pred. No.: 8.08e-74 Length: 1675
Score: 775.50 Matches: 185
Percent Similarity: 55.1% Conservative: 99
Best Local Similarity: 35.9% Mismatches: 181
Query Match: 30.1% Indels: 51
DB: 6 Gaps: 12
```

```
US-10-759-813-2 (1-500) x US-10-953-349-31583 (1-1675)
Qy 13 IleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGlnAen 32
Db 108 CTTGACATCGTGCCTCTCTCTCGCTTCGCCGTCTCTCTTGTATACATCTCTCGTAGG 167
Qy 33 Pro-----ProProGlyProTrp 38
Db 168 CCCGCCCTCTGCGCAGCGGCGATGAGGAGCGAGGCATCTTCCGCGTCACCGCGC 227
Qy 39 LysPheProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyGlu 58
Db 228 GGCCTCCCGCTGCTCGGC-----CACCTACATCTCTCGGCTCGTTCGCGCACCGG 278
Qy 59 ArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnVal 78
Db 279 GCCCTGCGGTCTCTGGCCGCCGACACACGCGCGTCTCTGCTCTCGGCTCGCGCGCTG 338
Qy 79 SerAlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAsp 98
Db 339 CCCGCGGTGTGTCTCTCGCGCGCGCGGAGGAGGTGATGAGGCGCGCACCTG 398
Qy 99 AlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAsp 118
Db 399 GCCTTCGCGAGCGCGCGCCCGAGCGCCATGCGCGACCGGCTCTCTACGGGCGC---GAC 455
Qy 119 ValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGlu 138
Db 456 GTAGGTTTCGGCCCTACGGCGAGTACTGGCGCAGCGCGCGCGCTGTGCTGTGTCCAC 515
Qy 139 PheLeuSerAlaLysLysValGlnSerSerArgLeuIleArg-GluGluGluMetGluAs 158
Db 516 CTCCTCAGCGCTCCGATCTCTCTCTCGCGCGCTCCGGGAGGAGGAGGCGCGCGC 575
Qy 158 pAlaIleThrPheLeuArgSerLysAla-----GlySerProValAsnIleThrLysI 176
Db 576 GCTGTGCGAGCGCGTCCCGGAGCGCGCGCGCGCGCGCTGTGGACCTGTGTGAGCT 635
Qy 176 IleTyrGlyIleIleIleSerIleMetIleArgThrSerValGlyAsnCysLys---- 194
Db 636 CCTCTCGCTAC-GCCACACCGTGTCTCGCGCGCGCGCTTCGGGAGCAGACGCGCG 694
Qy 195 -----GlnLysGluArgLeuLeuSerValAlaAspAlaValAaenGlu 208
Db 695 CGGCTGTATGAGGAAGCAACAGGAACGC-----GAGCTGAGGAAGGTGTTCAACGA 748
Qy 208 uAlaAlaThrSerPheGlyThrAla-----AspAlaPheProThrTrpLysLeuLe 225
Db 749 CTTCCAGGAGCTGCTCGGCACCGCGCCTCTGGGGAGCTCTCGCCCTGGCTGGGTGGT 808
Qy 225 uHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspI 245
```

Db 809 GGACGGCGTGAGGGGATGAGGGAAGATCAGACGGACGCTTCAAGCGCGCTCGACGGTGT 868
QY 245 eLeuGluGluLeuAAsnGluHisLysAlaAAsnLysProPhe----- 259
Db 869 GCTCGAAGAGGTGATCGGCGACCAACCGCCCGCGGTCAAGCCGGCCAGCAGACGGGAGA 928
QY 260 -----GluAlaAspAAsnLeuMetAAspValLeuLeuAAsnLeu---GlnLysAAsnG1 275
Db 929 TGACGGCGGCGCATCACAGGATTCGTGGACGTGTTGGACGTGAGCGATCTGATGA 988
QY 275 yAAsnValProValProValThrAAsnGluSerLysAlaSerValLeuGlnMetPheTh 295
Db 989 CGAAGCTGCGATCGCTCAGCAGCAGCCGAAATCAAGGCCATCATTTTGGACATGTCG 1048
QY 295 xAlaGlySerGluThrThrSerLysAlaThrGluTTPValMetAlaGluLeuMetLysAs 315
Db 1049 GCGCGGACGGACACAGCAGCAGCGAATGGAGTGGCCATGCGGAGGTCAACGCA 1108
QY 315 nProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLy 335
Db 1109 CCACAGACAGATCGTAAGCTCCAGCAGCAGCTCAGCGCGCGTCTCGGTGGCTCCGGCA 1168
QY 335 sVal---AAspGluSerArgPheHisAAspLeuLysPhePheLysLeuValValLysGluTh 354
Db 1169 CGTCATCACCGAGGACCATCGACAAAGCTGCACTACCTCAAGGCGGTGTCGAAGGAGAC 1228
QY 354 xLeuArgLeuHisProProVal---ValLeuLeuProArgGluCysArgGluThrThrAr 373
Db 1229 GCTACGCCCTGCACCCACCATCCCGCTCTCTGTCGCCCCGAGAGCCCAAGACGACGCCGA 1288
QY 373 gLeAAspGlyTyArgLysHisProAAsnThrArgLysValValAAsnAlaTTPAlaTleG1 393
Db 1289 GATACTCGGCCACACCGTCCCGCGCGGCGACCGCGGTGTCATCAACCGTGGGCGCTCGG 1348
QY 393 yArgAAspProAAsnThrTTPSerGluProGlyLysPheAAsnProGluArgPheLysAAspCy 413
Db 1349 CCGGGACCGCGCGCGTGGGAGCGCGCCGAGGAGTTCGTACCGAGAGGTTTCTCGACCG 1408
QY 413 sAlaTleAAspTyArgLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArg1 433
Db 1409 CCGGTGGACTACAGGGCGAGGACTTCCAGCTGATACCGTTCGTTGCTGGCGCGCGGG 1468
QY 433 eCysProGlyTleThrSerAlaTleThrAAsnLeuGluTyArgValLleAAsnLeuLeuTy 453
Db 1469 GTGCCCGGGTTCGGATTCGCCGCGCGCGCGGTTGAGATGGCGCTGGCGAGTTGATGA 1528
QY 453 rHisPheAAsnTTPGluLeuAlaAAspGlyLleThrProGlnThrLeuAAspMetThrGluAl 473
Db 1529 CCATTTTCGATCGGAGCGCGCGCGGCG-----TCGCTGGACATGCGCGAGGT 1576
QY 473 alleGlyGlyAlaLeuArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 487
Db 1577 GAACGGGCTCGCGGTGATCTCAAGTTCGCGCGCTGCGCGCTGTC 1619

RESULT 14

US-10-953-349-12597
; Sequence 12597, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12597
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-12597

Alignment Scores:
Pred. No.: 3,78e-72 Length: 1852
Score: 760.50 Matches: 170
Percent Similarity: 54.5% Conservative: 100
Best Local Similarity: 34.3% Mismatches: 196
Query Match: 29.5% Indels: 29
DB: 6 Gaps: 10
US-10-759-813-2 (1-500) x US-10-953-349-12597 (1-1852)
QY 18 ValLeuLeuValValValMetArgLeuTrpLysGlnAAsnPro---ProProGly 36
Db 119 ATCTCATATATCGTATAATAAAGTTTATAACACGCCCATCCCTCCCTCCACCGGGC 178
QY 37 ProTrpLysPheProLleGlyAAsnLeuProHisLeuLeuThrSerAAspLeuGly 56
Db 179 CCTAAACCATGGCCCATAGTGGGAAACTTGCCACACATG-----GGCCGGTGGCCG 229
QY 57 HisGluArgPheArgAlaLeuAlaGlnLleTyArgProValMetSerLeuGlnLleGly 76
Db 230 CACCACCTCCCTCGCGGGTGGCCCGCATCCACGGCCCGTTGATGACACCTCCGCTCGG 289
QY 77 GlnValSerAlaValLleSerSerAlaGluAlaLysGluValMetLysThrGln 96
Db 290 TTCGTCGATGTCGTGTTGCGGCTCGGCTCGGCGAACAGTCTTGAAGATTCAT 349
QY 97 AlaAAspAlaPheAlaGlnArgProLleValLeuAAspAlaGlnLleValPheTyArgAAsnArg 116
Db 350 GACTCTAATTTTCCAGCAGCGCGCCACCGAACCGCGCGCGAAATATATAGCTTATACTAC 409
QY 117 LysAAspValLeuPheAlaSerTyArgLysHisTTPArgGlnMetLysLysLysLysLys 136
Db 410 CAAGATCTGGTGTGGCCCTACCGCCACCGCTGGCGATTGCTCCGGAACATCACCTCC 469
QY 137 LeuGluPheLeuSerAlaLysLysValGlnSerSerArgLysLysLysLysLysLysLys 156
Db 470 GTTCACCTCTTCTCCGCGAAGCCCATGAATGAATTTAGACACTTGGCGTCAGGAAGGTA 529
QY 157 GluAAspAlaLleThrPheLeuArgSerLysAlaGlySerProValAAsnLleThrLysLys 176
Db 530 GCTAGATTGACATGCACTTGGCAAGTTTCAGACACAAAGCAGTGAATTTGGGACAAATTA 589
QY 177 IleTyArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 190
Db 590 TTGAATGATGACCCACCAACCGCATTTAGCGAGGCAATGATTCGACGAGAGTGTGTTAAC 649
QY 191 -----GlyAAsnCysLysGlnLys---GluArgLeuLeuSerValAlaAAspAlaVal 206
Db 650 GATGGCAATGGAGGGTGTGATCCTAGGGCTGATAGTATTAAGGCTATGGTGTAGGAGGTG 709
QY 207 AAsnGluAlaAlaThrSerPheGlyThrAlaAAspAlaPheProThrTrpLysLeuLeuHis 226
Db 710 ATGGGTGGCTGGAGTGTTCACATTTGATGCTTTATTCCTCTCTTGGAGTGGCTAGAC 769
QY 227 TyrlleLleGlyAlaGluSerLysProArgArgLeuHisGlnLleAAspAAspLleLeu 246
Db 770 ---CTTCAAGGGGTGCAAGCTAAGATAAATAATACACAAAGATTTGATGATTTTAA 826
QY 247 GluGluLeuLeuAAsnGluHis-----LysAlaAAsnLysPhePheGluAlaAAspAAsnLeu 264
Db 827 ACCAGCATATTATGGAGGACCAATAACAGTCTTCCCAAGAATGAGAAATCATAAGAATTC 886
QY 265 MetAAspValLeuAAsnLeuGln-----LysAAsnGlyAAsnValProValPro 280
Db 887 TTGAGTATCTTATGTCACTTAAAGATGTGAGGATGATCATGTGAAGAAC-----CAT 937
QY 281 ValThrAAsnGluSerLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 938 CTCACCTGATACCGAGATTAAAGCACTACTTCTGAACATGTTTACCGCAGCAACTGACACA 997
QY 301 ThrSerLysAlaThrGluTTPValMetAlaGluLeuMetLysAAsnProThrGluLeuArg 320
Db 998 TCATCAAGCACCCACAGATGGGCCAATATGCGGAATAATAAAAAAACCCACAAATTTTGGCC 1057

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:41:26 ; Search time 283 Seconds
(without alignments)
4958.769 Million cell updates/sec

Title: US-10-759-813-2

Perfect score: 2574

Sequence: 1 MSQKNLSPFSLISILLVLI.....KIDKLPIPVQVSLGNSIS 500

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=xlh

-Q/abs/ABSSWEB_spool/US10759813/runat_31052006_060309_2699/app_query.fasta_1

-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-HOST=abs02h -USER=US10759813 @CGN 1 1 204 @runat_31052006_060309_2699

-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq.*

2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq.*

3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq.*

4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq.*

5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq.*

6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq.*

7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq.*

8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq.*

9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq.*

10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1127.5	43.8	1781	3	US-09-499-302A-1
2	1119	43.5	1665	2	US-08-313-075A-29
3	1088	42.3	1666	2	US-07-912-900-24
4	1088	42.3	1666	2	US-08-285-309-24
5	1088	42.3	1666	2	US-08-502-046-24
6	993.5	38.6	1665	3	US-09-292-768-67
7	993.5	38.6	1665	3	US-09-292-768-69
8	992.5	38.6	1515	3	US-09-292-768-5

9	992.5	38.6	1665	3	US-08-881-784-8	Sequence 8, Appli
10	992.5	38.6	1665	3	US-09-292-768-3	Sequence 3, Appli
11	983.5	38.2	1762	3	US-08-881-784-5	Sequence 5, Appli
12	983.5	38.2	1762	3	US-09-292-768-1	Sequence 1, Appli
13	979.5	38.1	1762	3	US-09-292-768-65	Sequence 65, Appli
14	978.5	38.0	1762	3	US-09-292-768-63	Sequence 63, Appli
15	977.5	38.0	1691	3	US-08-948-564-3	Sequence 3, Appli
16	955.5	37.1	1762	3	US-09-172-339-5	Sequence 5, Appli
17	856.5	33.3	1929	3	US-09-380-420C-1	Sequence 1, Appli
18	856.5	33.3	1929	3	US-09-380-420C-1	Sequence 1, Appli
19	840	32.6	1838	3	US-08-948-564-1	Sequence 1, Appli
20	805.5	31.3	2013	3	US-09-615-192A-404	Sequence 404, App
21	801	31.1	1539	3	US-10-142-231-53	Sequence 53, Appli
22	801	31.1	1539	5	US-10-884-115-53	Sequence 53, Appli
23	776	30.1	1866	3	US-09-615-192A-103	Sequence 103, App
24	776	30.1	1866	3	US-09-169-789-103	Sequence 103, App
25	774.5	30.1	1883	3	US-08-991-677-3	Sequence 3, Appli
26	774	30.1	1657	3	US-08-948-564-11	Sequence 11, Appli
27	773	30.0	1737	3	US-09-142-108C-3	Sequence 3, Appli
28	767.5	29.8	1789	3	US-09-142-108C-3	Sequence 1, Appli
29	767	29.8	1824	3	US-09-142-108C-20	Sequence 20, Appli
30	755.5	29.4	1764	3	US-09-947-027-3	Sequence 3, Appli
31	747	29.0	1660	3	US-09-142-108C-16	Sequence 16, Appli
32	747	29.0	1838	2	US-09-091-432-1	Sequence 1, Appli
33	747	29.0	1838	3	US-09-387-663-1	Sequence 1, Appli
34	747	29.0	1838	3	US-09-214-139B-2	Sequence 2, Appli
35	744.5	28.9	1880	3	US-09-564-808-1	Sequence 1, Appli
36	743.5	28.9	1711	3	US-09-142-108C-5	Sequence 5, Appli
37	739.5	28.7	1748	3	US-09-142-108C-14	Sequence 14, Appli
38	739.5	28.7	1884	3	US-09-564-808-3	Sequence 3, Appli
39	733	28.5	1545	5	US-10-142-231-55	Sequence 55, Appli
40	733	28.5	1545	5	US-10-884-115-55	Sequence 55, Appli
41	731.5	28.4	1667	3	US-09-142-108C-22	Sequence 22, Appli
42	731.5	28.4	1835	3	US-09-564-808-5	Sequence 5, Appli
43	728.5	28.3	1730	3	US-09-672-785-3	Sequence 3, Appli
44	727	28.2	1663	5	US-10-146-374-1	Sequence 1, Appli
45	719.5	28.0	1737	3	US-09-126-420A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-499-302A-1

; Sequence 1, Application US/09499302A

; Patent No. 6369212

; GENERAL INFORMATION:

; APPLICANT: BOUNG-JUN, OH

; APPLICANT: MOON, KYUNG KO

; APPLICANT: YOUNG, SOON KIM

; TITLE OF INVENTION: A CYTOCHROME P450 GENE HIGHLY EXPRESSED IN THE

; FILE REFERENCE: 10324/P64443USO

; CURRENT APPLICATION NUMBER: US/09/499,302A

; CURRENT FILING DATE: 2000-02-07

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 1781

; TYPE: DNA

; ORGANISM: Capsicum annuum

US-09-499-302A-1

Alignment Scores:	1.15e-114	Length:	1781
Pred. No.:	1127.50	Matches:	214
Score:	67.3%	Conservative:	118
Best Local Similarity:	43.4%	Mismatches:	148
Query Match:	43.8%	Indels:	13
DB:	3	Gaps:	5

US-10-759-813-2 (1-500) x US-09-499-302A-1 (1-1781)

Qy 12 LeuileSerPheLeuLeuValLeuValMetArgLeuTrpLys 30

Db 25 TTAGTTGCGTTTGTCTCTCTCCAGCATCTTCTACTCAAAAAATGGAAAAACC 84
Qy 31 Gln-----AsnProProGlyProTrpLysPheProIleIleGlyAsnLeuProHis 48
Db 85 CAAGAGCTAACTTACCTCTCTGTCATGGAATACCTTTATTGGAGCCTACATCAC 144
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGly 68
Db 145 TTGGCAGTGGCAGGTCCTCTCATCATGCGCTAAAAAATCTAGCAAAATCTTTATGG 204
Qy 69 ProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGluAla 88
Db 205 CGGCTCATGACCTACGATCGGGGAAATTCCTACCGTCATCATTTGTCGCCCGGAATG 264
Qy 89 AlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProfileValLeuAsp 108
Db 265 GCGAGGAGTAGTAAACTCAGACCTCGCTTTCGCAACGAGCGCGAACTGTGGTG 324
Qy 109 AlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrp 128
Db 325 GCTGACATCGCTCATATGATAGTACGGATATAGCATTTTCTCCATATGTTGAATACTGG 384
Qy 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
Db 385 AGGCAGATTTCGTAATAATTTGCATACTCGAATCTCCTTAGTGCAGAGATGTCAAATCTTT 444
Qy 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
Db 445 AGCTCAATTCGCAGAGTAGCTGTGATGATGGTCTCATCATACGAACCATGCCAAT 504
Qy 169 SerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIleArgThr 188
Db 505 TTTCCCGTCAACCTTACAGCAAAATATTTGGTTTACAAGTTCCGTAACCTTGTAGATCA 564
Qy 189 SerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsn 207
Db 565 GCTCTGGGAAAATATGTGTCAGCAACAACTGATATTTTCATGAGGGAATAATA 624
Qy 208 GluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHisTyr 227
Db 625 TCATTGACAGGTGATTTAGTATTGCTGATTTTTCCTTACATGGAATAATGCTACATGAT 684
Qy 228 IleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIleLeuGlu 247
Db 685 GTTGGTGGTTCAAAAACTAGACTGCTGAAAGGCTCATGTAAAAATCGATGAGATTTGGAA 744
Qy 248 GluIleLeuAsnGluHisLysAlaAsnLys-----ProPhe 259
Db 745 CATGTAGTAGAATGAGCACAAACNGAATCGAGCGGATGGCCAAAGGTAAATGCGGNAATTT 804
Qy 260 GluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProVal 279
Db 805 GCGCGTGAAGATTGATGATGATGTTTGTCTAAGGGTTCGAGAAAGTCGAGAAATTCAAAT 864
Qy 280 ProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGlu 299
Db 865 TCCATCAGCGATGACAATATCAATCAATATTAGTGGACATGTTCTCCGCTGGATCTGAA 924
Qy 300 ThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeu 319
Db 925 ACGTATCGACAACATAATTTGGCATGATGATGATGATGATGATGATGATGATGATGATGAT 984
Qy 320 ArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSer 339
Db 985 GCAAGGCAACAGCTGAAGTGAGACAAAGTCTTCAAGGAAAAGAAAGGTTTCAACAAAT 1044
Qy 340 ArgPheHisAspLysPhePheLysLeuValValLysGluThrLeuArgLeuHisPro 359
Db 1045 GATCTTTGATGAGTTGAAGTACTTGAAGTTAGTAAATCAAGAAACTCTAAGGATGACCTT 1104
Qy 360 ProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGlu 378

Db 1105 CCAATTCCTCTATTAGTCCCTAGAGATGTATGAGGATACAAAGATTGACGGGTACAAAT 1164
Qy 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db 1165 ATACCTTTTAAACTCGAGTCATAGTTAATGATGGCAATTGGACGAGATCCTGAAAGT 1224
Qy 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
Db 1225 TGGGATGACCTGAAAGCTTTTCCCAAGAGATTCGAGATAGTTCTGTGACATTTCTTT 1284
Qy 419 GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db 1285 GGAACCATCATCAATTTATTCATTTGGTGGGGAAGAGATTGCTCTGGAATGCTT 1344
Qy 439 SerAlaIleThrAsnLeuGluTyrValIleIleLeuLeuLeuLysHisPheAsnTrpGlu 458
Db 1345 TTTGTTTGTAGCAATGTTGGACCAACCATTAGCTCAATTTATCTACTTCGATCGGAAA 1404
Qy 459 LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeu 478
Db 1405 CTCCCTAATGACAAAGTCACGAAATTTGGACATGAGGATCACCTGGAATTTCTGCA 1464
Qy 479 ArgLysLysIleAspLeuLysLeuIleProfileProTyr 491
Db 1465 ACAGAAAGGATGATCTTGTGTTGATTGCCACCCCTTAT 1503

RESULT 2

US-08-313-075A-29
; Sequence 29, Application US/08313075A
; Patent No. 5639870
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwina C.
; APPLICANT: Tanaka, Yoshikazu
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,075A
; FILING DATE: 30-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 1538/92
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 6698/93
; FILING DATE: 07-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU93/00127
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9433
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs

; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3..1432
 ; US-08-313-075A-29

Alignment Scores:

Pred. No.: 9,06e-114 Length: 1665
 Score: 1119.00 Matches: 223
 Percent Similarity: 64.1% Conservative: 96
 Best Local Similarity: 44.8% Mismatches: 163
 Query Match: 43.5% Indels: 17
 DB: 2 Gaps: 5

US-10-759-813-2 (1-500) x US-08-313-075A-29 (1-1665)

```

QY 12 LeuileSerPheLeuLeuValLeuValValMetArgLeuTrpLysGln 31
Db 15 TTGGTTTCCTTCCTTATGTTATTTCCCTCATTTTCTAAGAAATGGAGAATCC 74
QY 32 Asn-----ProProGlyProTrpLysPheProIleleGlyAsnLeu 46
Db 75 AATTGTCAAAACCAAAATTTGCCCTCCAGGCCCATGGAAAGTACCTTTCTTGAAGCTTG 134
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuGlnIle 66
Db 135 CTTTCATATGGTA-----GGTGGACTTCCACACCATGTCTTAGAGATTTCAGCAAAAAA 188
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 189 TATGGACCAATATGACCTTCACTAGGTAAATTTTCGCCGTTGTAGTACTTCTCCT 248
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 249 GAGATGGCAAGAAAGTACTAAAAAATCATGACCTTGCATTTGCATATAGGCCTAAACTT 308
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 309 CTAGGCATTGAGATTGTCTGCTATAATAGTTTCAGACATTGCCCTTTTCCCGGTATGGTAT 368
QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 369 TACTGGAGCAATGGTAAATTTTGTATTGGAAGTGTCTAGTSCCAAAATGTCGG 428
QY 147 SerSerArgLeuLeuArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 429 TCATTTTAATTCGATTAGACGAGATGAATACTTCTTATGATCGATTTTTCGCGATCATCT 488
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
Db 489 TCTGGTAACCGAGTTAATATAACAGAAAGATCTTTTCATTCACAGCTCTATGATTTGT 548
QY 187 ArgThrSerValGly---AsnCysLysGlnLysGluArgLeuSerValAlaAspAla 205
Db 549 AGATCAGTATTTGGGAAGAGATAAAGAGAGACGAATGTATACGATGTGAAAAA 608
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 609 ATGACAGGCTTAATAGATGGTTCGATGTGGCTGACATATTCCTTCGTTGAGGTTCTT 668
QY 226 HisTyrIleleGlyAlaGluSerLysProArgArgLeuHisGlnGluLeuAspIle 245
Db 669 CATGTACTAATCGGTATGAGGGTAAATTTATGGATGTTTCATCGTAAGGTAGTACTATT 728
QY 246 LeuGluGluLeuLeuAsnGluHis-----LysAlaLeuLysPro 258
Db 729 GTTGAGGAATCATGATGAGCAAAAGAAACTCTTCGAACTGGCGAAGCAATGTGTAA 788
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValPro 278

```

```

Db 789 GTGGGAGGAGAGAGATTAAATTTGATGTTATCTTAAGACTTTAAGGAGAGGAGACCTTCAA 848
QY 279 valProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 849 CTTCCAATCACAATGACAACATCAAGGCCATTTTAAATGACATGTTTGTGCGGGAACA 908
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 909 GAAACTTCATCAACAACAATTAACCTGGGCCATGGTAGACTGATGAAAAATCAAGTGTA 968
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 969 TTGCGCAAAAGCTCAAGCAGAGGTAAAGAGAGTCTTCAAGGGGAAGAAACTTTTCGATGA 1028
QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1029 GATGATATCAGGAGCTGAATTACCTTAAGTTAGTTCATTAGAGAAACTTTTAAGACTCCAC 1088
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1089 CCTCCACTTCCACTTTTGTCTTCCAGAGATGTCCGAGAGAAACAGAAATTAATGGCTAC 1148
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1149 ACTATTCTCTTAAATACCAAAAGTCATAGTTAATGTTGGCTATTGCAAGAGATCCAAAA 1208
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1209 TATGGGATGATCGAAGAGCTTTAAGCCTGAGAGATTTGAACATAACTCTTTGCAATTTT 1268
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1269 GCTGGCAATAATTTGAATATCTCTTTTGTGTAGTGGAGAGGATTTGCCCGGNATA 1328
QY 438 ThrSerAlaIleThrAsnLeuGluTrpValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1329 TCATTTGGTTAGCTAATGTTATCATCCATGGCTCAATTTGTTGTTATCATTTTCGATTGG 1388
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrCluAlaIleGlyGlyAla 477
Db 1389 AGACTTCTCTACTGCTGGGTGACCCCAATGACTTTGAATTCAGTAGTTA-GCTGGAGTAACT 1447
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeu 495
Db 1448 ACTGGTAGAAGAAAGAGACCTTTACTTGATTTTCACCTCTTATTCACCTTCTCTA 1501

```

RESULT 3

US-07-912-900-24
 ; Sequence 24, Application US/07912900
 ; Patent No. 5349125

GENERAL INFORMATION:

; APPLICANT: Holton, Timothy A.
 ; APPLICANT: Cornish, Edwina C.
 ; APPLICANT: Kovacic, Filipa
 ; APPLICANT: Tanaka, Yoshikazu
 ; APPLICANT: Lester, Diane R.
 ; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
 ; NUMBER OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

; ADDRESSER: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/912,900

```

; FILING DATE: 19920713
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8613
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-912-900-24

Alignment Scores:
Pred. No.: 2,48e-110 Length: 1666
Score: 1088.00 Matches: 221
Percent Similarity: 62.9% Conservative: 93
Best Local Similarity: 44.3% Mismatches: 168
Query Match: 42.3% Indels: 18
DB: 2 Gaps: 5

US-10-759-813-2 (1-500) x US-07-912-900-24 (1-1666)

Qy 12 Leu1eSerPheLeuLeuValLeuValVal-MetArgLeuTrpLysLysG1 31
Db 15 TTGGTTTCCTTCTCTTATTGTTATTTCCCTCATTTTCATTAGAAATGGAAGAAATC 74
Qy 31 nAen-----ProProGlyProTrpLysPheProIleleGlyAsnLe 46
Db 75 CAATTGTCAAACCAAAAAATTGCCCTCCAGGCCCATGGAAAGTAGACCTTTTCTTGGAGCTT 134
Qy 46 uProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnI 66
Db 135 GCTTCATATGGTA-----GGTGGACTTCCACACCATGTCTCTTAGAGATTTAGCCAAAAA 188
Qy 66 eTrpGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAl 86
Db 189 ATATGAGACCAANTTATGCACCTTCACTAGGTAAATTTCTGCCGTGTAGTTACTTCTCC 248
Qy 86 aGluAlaLalLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProfileVa 106
Db 249 TGAGATGGCAAGAAAAGTACTAAAAAATCATGACCTTGCATTTGCATATAGGCCTAAACT 308
Qy 106 lleuaspAlaGlnIleValPheTyrrAsnArgLysAspValLeuPheAlaSerTyrrGlyAs 126
Db 309 TCTAGGCATTGAGATTGTCTGCTATTAATAGTTTCAGACATTTGCCCTTATGGTGA 368
Qy 126 phiSTrpArgGlnMetLysLysIleTrpIleLeuGluPhLeuSerAlaLysLysValG1 146
Db 369 TTACTGAGGCAAAATGGCTAAATTTGTGTATTGGAAGTCTTAGTGCCTAAATGTCGCC 428
Qy 146 nSerSerArgLeulleAArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerIy 166
Db 429 GTCATTTAACTCGATTAGACGAGATGAATTAATCTTTTATGATCGATTTTTTTCGCATCATC 488
Qy 166 salAGlySerProValAsnIleThrLysIleIleTyrrGlyIleIleIleSerIleMetI1 186
Db 489 TCTCGTGAAGCCAGGTAAATATAACAGAAAGATCTTTTTCATTCAAGCTCTATGATTTG 548
Qy 186 eArgThrSerValGlyI--AsnCysLysGlnLysGluArgLeuLeuSerValAlaAspAl 205
Db 549 TAGATCAGTATTTGGGAAAAGAAATAAAGAGAAAGACGACGATGTATACGATGTGAAAAA 608
Qy 205 aValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLe 225
Db 609 AATGACAGGCTTAAATAGATGGGTTCCGATGTGGCTGACATATTCCTTCGTGTGAGGTTCT 668

```


APPLICANT: Kovacic, Filippa
APPLICANT: Tanaka, Yoshikazu
APPLICANT: Lester, Diane R.
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
TITLE OF INVENTION: HYDROXYLASE AND USES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/502,046
FILING DATE: 14-JUL-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/285,309

FILING DATE: 03-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8633Z

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 1666 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-502-046-24

Alignment Scores:
Pred. No.: 2,48e-110 Length: 1666
Score: 1088.00 Matches: 221
Percent Similarity: 62.9% Conservative: 93
Best Local Similarity: 44.3% Mismatches: 168
Query Match: 42.3% Indels: 18
DB: 2 Gaps: 5

US-10-759-813-2 (1-500) x US-08-502-046-24 (1-1666)

QY 12 LeuileSerPheLeuLeuValLeuValVal-MetArgLeuTyrLysLysG1 31
DB 15 TTGGTTCTCTTCCTCTATTGTAATTTCCCTCATTTTCATTAGAAATGGAAGATC 74
QY 31 nAen-----ProProGlyProTyrLysPheProilelleGlyAsnLe 46
DB 75 CAATTGTCAACCAAAAAATTCCTCCAGGCCCATGGAAAGTACCTTTTCTTGGAGCTT 134
QY 46 uProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln1 66
DB 135 GCTTCATATGGTA-----GGTGGACTCCACACCATGCTCTTAGAGATTAGCCAAAA 188
QY 66 eTyrGlyProValMetSerLeuGlnleGlyGlnValSerAlaValValIleSerAl 86
DB 189 ATATGACCAATATATGACCTTCACCTAGTAAATTTCTGCGTTGATGTTACTTCTCC 248
QY 86 aGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVa 106
DB 249 TGAGATGGCAAGAAAGTACTAAAAAATCATGACCTTGCAATTTGCATATAGGCTAAACT 308
QY 106 lleuAspAlaGlnleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAs 126

DB 309 TCTAGGCATTGAGATTGTCTCTATAATAGTTCAGACATTCGCTTTTCCCGTATGGTGA 368
QY 126 pHisTyrArgGlnMetLysLysIleTyrPileLeuGluPheLeuSerAlaLysLysValG1 146
DB 369 TTACTGGAGGCAATGCGTAAATTTGTGTATTGGAAGTCTTAGTGCCAAAAATGTCGG 428
QY 146 nSerSerArgLeuLeuArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerly 166
DB 429 GTCATTTAACTCGATTAGACGAGATGAATATCTTCTTATGATCGAATTTTTCGCGATC 488
QY 166 salAglySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetI1 186
DB 489 TCTCGTAAAGCAGTTATATATAACAGAAAGATCTTTTCATTCACAACTCTATGATTG 548
QY 186 eArgThrSerValGly---AenCysLysGlnLysGluArgLeuLeuSerValAlaAspAl 205
DB 549 TAGATCAGTATTTGGGAAAGAAATAAGAGAGAGCAATGATATACGACATGTGAAAAA 608
QY 205 aValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPheProThrTyrLysLeuLe 225
DB 609 AATGACAGGCTTAATAGATGGGTTTCGATGTGGCTGACATATTCCTTCGTGAGGTTCT 668
QY 225 uHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnIleAspAspI1 245
DB 669 TCATGTACTAATCGGTATGAAGGGTAAATTTATGGATGTTTCATCGTAAGGTAGATCTAT 728
QY 245 eLeuGluGluLeuLeuAsnGluHis-----LysAlaAsnLysPr 258
DB 729 TGTGAGGAAGCTATGAATGAGCACAAAGAAACTCTTCGAACTGGCAAGACCACTGGTGA 788
QY 258 oPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPr 278
DB 789 AGTGGGAGGAGAAAGATTTAATGTATGTTGCTTAAGACTTAAGGAAGAGGAGACCTTCA 848
QY 278 oValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySe 298
DB 849 ACTTCCAATCAAAATGACACACTAAAGCCATTTTAAATGACATGTTTGTGCGCGGAAC 908
QY 298 rGluThrThrSerLysAlaThrGluTyrValMetAlaGluLeuMetLysAsnProThrG1 318
DB 909 AGAAACTTCATCAACAACAATTAACCTGGGCCCTAGTAGAATCTGATGAAAAATCCACGT 968
QY 318 uLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspG1 338
DB 969 ATTCCGAAAGCTCAAGCAGAGGTAAAGAGAGTCTTCAAGGGAAGAAACTTTTCGATGA 1028
QY 338 uSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHi 358
DB 1029 AGATGATATCGAGGAGCTGAATTAACCTTAAGTTAGTTCATTAGAGAAACTTTAAGATCTCA 1088
QY 358 sProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTy 377
DB 1089 CCTTCCACTTCCACTTTTCTTCCAAAGAGAAATGTCGAGAGAAACAGAAATAATATGGCTA 1148
QY 377 rGluIleHisProAsnThrArgIleValValAsnAlaTyrAlaIleGlyArgAspProAs 397
DB 1149 CACTATTCTTTAAATACCAAGTCATAGTTAATGTTGGGCTATTGGAGAGATCCAAA 1208
QY 397 nThrTyrSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTy 417
DB 1209 ATATTGGGATGATGAGAAAGCTTTAAGCCTGAGAGATTTGAACATAACTCTTTGAATTT 1268
QY 417 rLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyI1 437
DB 1269 TGCTGGCAATAATTTTGAATATCTTCTTTTGGTATGTAAGAGAGATTTGCCCGGAAT 1328
QY 437 eThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTr 457
DB 1329 ATCATTTGGTTAGCTAATGTATCATCCATGGCTCAATTTGTTGTTATCATTTTCATG 1388
QY 457 pGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAl 477

QY 481 LysileAspLeuLysLeuLeuProLeuProTyrGlnValSer 494
 Db 1474 AAGAAATCTTCTACTCGTCCACACCCCTACGATCCTTCC 1515

RESULT 7

US-09-292-768-69
 ; Sequence 69, Application US/09292768
 ; Patent No. 6194185
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney B
 ; APPLICANT: Lupien, Shari L
 ; APPLICANT: Karp, Frank
 ; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
 ; TITLE OF INVENTION: LIMONENE HYDROXYLASES
 ; FILE REFERENCE: wslr13463
 ; CURRENT APPLICATION NUMBER: US/09/292,768
 ; EARLIER FILING DATE: 1999-04-14
 ; EARLIER FILING DATE: 08/891,784
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 69
 ; LENGTH: 1665
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
 ; OTHER INFORMATION: limonene-3-hydroxylase
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1665)
 ; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
 ; OTHER INFORMATION: limonene-3-hydroxylase protein variant
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (19)..(1518)
 ; US-09-292-768-69

Alignment Scores:

Pred. No.: 7,37e-100 Length: 1665
 Score: 993.50 Matches: 205
 Percent Similarity: 61.7% Conservative: 100
 Best Local Similarity: 41.5% Mismatches: 178
 Query Match: 38.6% Indels: 11
 DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-09-292-768-69 (1-1665)

QY 10 SerileLeuileSerPheLeuLeuValLeuileLeuValValMetArgLeuTrpLys 29
 Db 40 GCGATTATAATCTTGTAGTAACCTACACCATATCCTCCTAATAATCAAGCAATGGCGA 99
 QY 30 LysGlnAsnPro-----ProProGlyProTrpLysPheProIleileGlyAsn 45
 Db 100 AAACCCGAAACCCCAAGAGAACCTGCTCCGGGCGCCGGAAGCTGCCGTGATCGGGCAC 159
 QY 46 LeuProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
 Db 160 CTC---CAC---CTCTATGGGGGAAGCTGCCGACGACGCGCTGGCCAGCTGGCGAAG 213
 QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
 Db 214 CAGTACGGCCAGTGGCGCACGTGCAGCTCGCGAGGTGTCTCCGCTGCTGCTCGTCC 273
 QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
 Db 274 CCGAGGCCACGAGGAGCGGATGAAGCTGGTGGACCCCGGCTGGCGGACCGGTTCGAG 333
 QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
 Db 334 AGCATCGGACGAGATCATGTGGTACGACACGACGACATCATCTTCAGCCCTACAGC 393

QY 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
 Db 394 GTGCATCGGGCCAGATCGCGAAGATCTGCGTCTCCGAGCTCCTCAGCGCCGCGCAACGTC 453
 QY 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSer 165
 Db 454 CGTCTCTTCGGCTTCATCAGCAGCAGAGGTGTCCCGCTCCTCGGCCACCTTCGCTCC 513
 QY 166 Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
 Db 514 TCGGCGCGCGGGAGCGCTGACCTCAGCAGCGGATAGCGACGCTGACGTGCTCC 573
 QY 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerVal 202
 Db 574 ATCATCTGCGGCGCGGTTCGGGAGCGTATCATCGGACCGACGAGGCTGTGTGAGCTG 633
 QY 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
 Db 634 GTGAAGGAGCGCTCAGCATGCGTCCGGTTCGAGCTCGCGACATGTTCCCTCTCC 693
 QY 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluLe 242
 Db 694 AAGCTCTCAACTGCTCTGCTGGAACAAGAGCAAGCTGTGGAGGATGCGCGCGCTC 753
 QY 243 AspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys-----ProPheGlu 260
 Db 754 GACGCCATCTCGAGGCGCATCGTGAGCAGACCAAGCTCAAGAGCGGCGCGAGTTGGC 813
 QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValProValPro 280
 Db 814 GCGGAGGACATATTGAGCTACTCTTTAGGATGCAGAGGATAGCAGATCAAGATCCCC 873
 QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
 Db 874 ATCACCACCAACGCCCATCAAGGCTTCATCTCGACACGTTCTCAGCGGGGACCGAGACA 933
 QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
 Db 934 TCATCAACCAACCCCTGTGGTGTGATGGGAGCTGTAGGATCCAGAGGTGATGGCG 993
 QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
 Db 994 AAAGCGCAGCGGAGGTGAGAGCGCGCTGAAGGGGAAGAGCGAGCTGGGACGTGGACGAC 1053
 QY 341 PheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisProPro 360
 Db 1054 GTGCAGGAGCTTAAGTACATGAATCGGTGTGTAAGAGGACGATGAGGATGACCCCTCG 1113
 QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
 Db 1114 ATCCCGTTGATCCCGAGATCATGCGAGAGAAGATCGGAGGTCAAGGGTACAGATTCCG 1173
 QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
 Db 1174 AATAGGCCAGATCATGATCAACCTGTTGCTCCATGGGTAGGATCTCTCTACTGGAA 1233
 QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
 Db 1234 AAACCCGAGACCTTTTGGCCCGAAGGTTTGACCAAGTCTCGAGGGAATTCATGGGAAC 1293
 QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
 Db 1294 GATTTCCGAGTTTATCCCATTTGGAGCTGGAAGAAGATCTGCCCGGTTTGAATTCGGG 1353
 QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
 Db 1354 TTGGCAATGTTGAGGTCCCATTTGGCACAGCTTCTTACCCTTCGACTCGAAGTGGCG 1413
 QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLys 480
 Db 1414 GAAGGAATGAACCTTCCGATATGACATGTCGTGAGGCAAGAGGCTTACCGGAATAAGA 1473

QY 481 LysileAspLeuLysLeuLeuProileProTyrGlnValSer 494
Db 1474 AAGAACAACTCTTCTACTCGTTCCACACACCTACGATCCTTCC 1515

RESULT 8

US-09-292-768-5
; Sequence 5, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wsu13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1494)
US-09-292-768-5

Alignment Scores:

Pred. No.: 8,21e-100 Length: 1515
Score: 992.50 Matches: 206
Percent Similarity: 61.5% Conservative: 98
Best Local Similarity: 41.7% Mismatches: 177
Query Match: 38.6% Indels: 13
DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-09-292-768-5 (1-1515)

QY 12 LeuileSerPheLeuLeuValLeuLeuValValMetArgLeu----- 27
Db 16 CTTTGTGCGCGCTTATAATCTCTGTAGTAACCTACACCATATCCCTCTTAATCAACCAA 75
QY 28 TrpLysLysGlnAsnPro-----ProProGlyProTrpLysPheProIleIle 43
Db 76 TGGCGAAACCCGAAACCCCAAGGAAAGTTCCCGCGGCGCGCGGAGGTGCGGTGATC 135
QY 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db 136 GGGCACCTC---CAC---CTCTGTGGGGAGCTCCCGCAGCAGCAGCGGTGGCCAGCGTG 189
QY 64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIle 83
Db 190 GCGAAGGAGTACGGCCCGCGCGCCAGTGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 249
QY 84 SerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 250 TCGTCGGGAGCGGCGAAGAGGCGGATGAAGCTGTAGACCCCGCGCGCGCGCGCGCGG 309
QY 104 ProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
Db 310 TTCGAGAGCATCGGACGAGGATCATGTGTAGCAACAGCAGGACATCATCTTCAGCCCC 369
QY 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db 370 TACAGGAGCACTGGCGCCGAGATCGCAAGATCTGGTCTCCGAGCTCTCTCTCTCCCGC 429
QY 144 LysValGlnSerSerArgIleAlaArgGluGluGluMetGluAspAlaIleThrPheLeu 163
Db 430 AACGTCGGCTCTTCGGCTTACCGGACGAGGAGTGTGCGCGCTCTCCGCCACCTC 489
QY 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleIleIleIleIleSer 183

Db 490 CGCTCTCGCGAGCGCGCGCTGGACATCAGCGAGGATAGAGCGCTGACGTGCTCC 549
QY 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerVal 202
Db 550 ATCATCTGCGAGCGCGCTTGGGAGCGTATCAGGACAAACCGCGAGCTGGTGGGCTG 609
QY 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
Db 610 GTCAGGAGCGCTCAGCATGGCTCGGGTTTGAGCTCGCCGACATGTTCCCTCTCCTCC 669
QY 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgLeuGlnHisGlnIle 242
Db 670 AAGCTCTCAACCTCTCTGCTGAACAAGCAAGCAAGCTCTGGAGGATGCCCGCGGCTC 729
QY 243 AspAspIleLeuGluGluLeuLeuAsnGluHisLysAlaAsnLys-----ProPheGlu 260
Db 730 GACACCATCTCAGGCGCATCTGTCGACGACACAACTTCAAGAAGAGCGCGAGTTCGGC 789
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
Db 790 GCGGAGGACATCATCGACGTCTCTTCAGGATGCAAGAAGCCACCCAGATCAAGTCCCC 849
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 850 ATCACCACCACTCTCATCAAGCTTCTTCGATACGTTCTCAGCAGGAGTCTGAGACA 909
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
Db 910 TCCTCAACCAACCCCTATGGGTGCTGGCGGAGCTGATGAGGAACCCCGCAGTGTGGCG 969
QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
Db 970 AAAGCGCAGCGGAGGTGAGCGCGCATGTAAGAGAGAGCAAGCAAGCTGGGACGTGGATG 1029
QY 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
Db 1030 GTGCAAGAGCTTAAGTACATCAATCGGTGGTGGAGGACGATGAGGATGCACTCCG 1089
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
Db 1090 ATCCCGTTGATCCGAGATCATGTCAGAGAAGAATGCGTGTAAACGGGTATACGATTCCG 1149
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db 1150 AACAGCGCAGAAATCATGATCAACGTCTGTCATGGCGAGGAATCCTCTTACTGGGAA 1209
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
Db 1210 AACCCGATACCTTTTGGCCCGAAGTTTGACCAAGTTTCAAGAGATTTCATGGGAAAT 1269
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
Db 1270 GATTTCGAGTTCGTCGCGTTCGAGCGGGAAGAATCTGCCCGCTTGAATTCGGG 1329
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1330 TCGGCAAGGTTGAGTTCATTTGGCGCAGCTTCTTACCATTCGACTGGAAGTTGGCG 1389
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1390 GAAGGATGAACCTTCTGATATGACATGCTGAGGCGGAGGCTTACCGGAATACTA 1449
QY 481 LysIleAspLeuLeuLeuProIleProTyrGlnValSer 494
Db 1450 AAGAACAACTCTCTTCTTGTTCCTCCACACCTACGATCCTCTCC 1491

RESULT 9

US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.

APPLICANT: Lupien, Shari L.
 APPLICANT: Karp, Frank
 TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
 THE PRODUCTION OF LIMONENE HYDROXYLASES
 NUMBER OF SEQUENCES: 58
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
 ADDRESS: PLLC
 STREET: 1420 Fifth Avenue, Suite 2800
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/881,784
 FILING DATE:

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Shelton, Dennis K.
 REGISTRATION NUMBER: 26,997
 REFERENCE/DOCKET NUMBER: WSUR19777
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 224-0718
 TELEFAX: (206) 224-0779

INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1665 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: Mentha x piperita
 IMMEDIATE SOURCE:
 CLONE: pPWL7
 US-08-881-784-8

Alignment Scores:
 Pred. No.: 9.51e-100 Length: 1665
 Score: 992.50 Matches: 205
 Percent Similarity: 61.5% Conservative: 99
 Best Local Similarity: 41.5% Mismatches: 179
 Query Match: 38.6% Indels: 11
 DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-08-881-784-8 (1-1665)

QY 10 SerIleLeuIleSerPheLeuLeuValValLeuLeuValValMetArgLeuTrpLys 29
 Db 40 GCATTATATCTCTAGTAACCTACACATATCTCTCTAATATCAAGCAATGGCGA 99
 QY 30 LysGlnAsnPro-----ProProGlyProTrpLysPheProIleLeuGlyAen 45
 Db 100 AAACCGAAACCCCAAGAGAACCTGCTCCGGGCGCCCGCGAAGTGCCTGATCGGGCAC 159
 QY 46 LeuProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
 Db 160 CTC---CAC---CTCCTATGGGGAGAGCTGCCGACGACGCGCTGGCCAGCGTGGCGAAG 213
 QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
 Db 214 CAGTAGCGCCCATGGCGGCAGCTGCGGAGGTGTCTCCGCTGCTGCTCTCTGCTCC 273
 QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
 Db 274 CGCGAGGCCACGAAGGCGGATGAAGCTGGTGACCCCGCCTGCGCGGACCGGTTTCGAG 333
 QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125

Db 334 AGCATCGGGACGAAGATCATGTGTACGACAAACGACGACATCACTTCAGCCCTACAGC 393
 QY 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
 Db 394 GTGCACCTGGCGCCAGATGCGGAAGATCTGCGTCTCCGAGCTCTCAGCGCCGCAACGTC 453
 QY 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleTrpPheLeuArgSer 165
 Db 454 CGCTCTTCGGCTTCATCAGGACGAGAGTGTCCCGCTCTCGGCCACCTCCGCTCC 513
 QY 166 Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
 Db 514 TCGGCGCGGGGGAGCGCTGAGACCTCAGGAGCGGATAGCAGCCTGACGTGCTCC 573
 QY 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerVal 202
 Db 574 ATCATCTGCAGGCGCGCTTCGGGAGCGTATCAGGACACGAGGAGCTGGTGAGGTG 633
 QY 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
 Db 634 GTGAAGGACGCCCTCAGCATGCGGTTCGGGTTCGAGCTCGCGCATGTTCCTCTCTCC 693
 QY 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIle 242
 Db 694 AAGCTCTCAACTTCTGCTCTGCTGGNACAAGAGCAAGCTGTGGAGGATGCGCGCGCTC 753
 QY 243 AspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys-----ProPheGlu 260
 Db 754 GAGCCATCTCTCGAGGCCATCGTGAGGAGCACAAAGCTCAAGAGAGCGCGAGTTTGGC 813
 QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
 Db 814 GCGGAGGACATTATTGACGTACTCTTTAGGATGCAAGAGGATAGCCAGATCAAGATCC 873
 QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
 Db 874 ATCACCACCAAGCCATCAAGCCCTTCATCTTCGACAGCTTCTCAGCGGGAGCCAGACA 933
 QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
 Db 934 TCATCAACCAACCACCTGTGGGTGATGCGGAGCTGATGAGGAATCCAGAGGTGATGGCG 993
 QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
 Db 994 AAAGCCGAGCGGAGGTGAGAGCGCGCTGAAGGGGGAAGACCGACTGGGACCTGGACGAC 1053
 QY 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
 Db 1054 GTGCAGGAGCTTAAGTACATGAATCGTGTGTGAAGGACGATGAGGATGACCCCTCCG 1113
 QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
 Db 1114 ATCCCGTTGATCCCGAGATCATGCAGAGAAGAATGCGAGTCAACGGGTACAGATTCGG 1173
 QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
 Db 1174 AATAAGGCCAGAAATCATGATCAACGTGTGTCCATGGGTAGGAATCTCTCTACTGGGA 1233
 QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
 Db 1234 AAACCCGAGACCTTTTGGCCGAAAGTTTGACCAAGTCTCGAGGATTCATGGGAAAC 1293
 QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
 Db 1294 GATTTTCAGTTCCATCCCATTTGGAGCTGGAAGAAGAAATCTGCCCGGTTTGAATTCGGG 1353
 QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
 Db 1354 TTGGCAAAATGTTGAGGTCCCATTTGGCAGAGCTCTTTACACTTCGACTTGGAGTTGGCG 1413
 QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLys 480

Db 1414 GAAGGAATGAACCTTTCCGATATGGACATGCTCTGAGGCAGAGGCTTACCGGAATAAGA 1473
QY 481 LysIleAspLeuLysLeuLeuProIleProTyrGlnValSer 494
Db 1474 AAGACAATCTTCTACTCGTTCCACACACCTAGATCCTTCC 1515

RESULT 10

US-09-292-768-3
; Sequence 3, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wu013463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-3

Alignment Scores:

Pred. No.: 9,51e-100 Length: 1665
Score: 92.50 Matches: 205
Percent Similarity: 61.5% Conservative: 99
Best Local Similarity: 41.5% Mismatches: 179
Query Match: 38.6% Indels: 11
DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-09-292-768-3 (1-1665)

QY 10 SerIleLeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLys 29
Db 40 GCGATTATAATCTTGATGTAACCTACACCATATCCCTCTTAATCAAGCAATGGCA 99
QY 30 LysGlnAsnPro-----ProProGlyProTrpTrpLysPheProIleIleGlyAsn 45
Db 100 AAACCGAAACCCCAAGAGAACCTGCTCCGGGCGCCGCGGAAGCTGCGCTGATCGGGCAC 159
QY 46 LeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
Db 160 CTC---CAC---CTCTATGGGGAGCTGCCGAGCAGCGCTGGCCAGCGTGGCGAAG 213
QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
Db 214 CAGTACGGCCAGTGGCGACGCTGAGCTGCGGAGGTGTTCTCGTGTGCTCGTCC 273
QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
Db 274 CGCGAGGCCACGAAGAGCGGATGAAGCTGGTGGACCCCGCTGGCGGACCGGTTTCGAG 333
QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
Db 334 AGCATCGGACGAAGATCATGTGTACGACAAACGACGACATCATCTTCAGCCCCCTACAGC 393
QY 126 AspHisTrpArgGlnMetLysIleTrpIleLeuGluPheIleuSerAlaLysLysVal 145
Db 394 GTGCACTGGCGGCAGATGCGGAAGATCTCGCTCTCCGAGCTCTCAGCGCCCGCAACGTC 453
QY 146 GlnSerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSer 165
Db 454 CGCTCTTCGGTTCATCAGGAGGACGAGGTGTCCCGCTCTCCGCGCCACCTCCGCTCC 513

RESULT 11

US-08-881-784-5
; Sequence 5, Application US/08881784
; Patent No. 6083731

QY 166 Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 514 TCGCGCGCGCGGGAGCGCGTGGACCTCAGAGCGGATAGCGACGCTGTCGCTCC 573
QY 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuSerVal 202
Db 574 ATCATCTGCAGGCGCGCTTCGGAGCGTGATCAGGACCAACGAGGAGCTGTGGAGCTG 633
QY 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
Db 634 GTGAAGGACGCCCTCAGCATGGCTCCGGGTTCGAGCTCGCCGACATGTTCCCTCTCC 693
QY 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIle 242
Db 694 AAGCTCTCTCAACTTGTCTGCTGGAAACAGAGCAAGCTGTGGAGATGCCGCCCGCTC 753
QY 243 AspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys-----ProPheGlu 260
Db 754 GAGCCATCTTCGAGGCCATCTGGAGGAGCACAGCTCAAGAGAGCGCGGAGTTGGC 813
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
Db 814 GCGGAGGACATTATTGACGTACTCTTTAGGATGCAGAGGATAGCCAGATCAAGATCCCC 873
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 874 ATCACCACCAACGCCATCAAGCCTTCATCTTCGACACCGTTCACGCGGAGGACAGACA 933
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
Db 934 TCATCAACCCACCACCTGTGGGTGATGCGGAGCTGATCAGGAATCAGAGGTGATGGCG 993
QY 321 LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
Db 994 AAAGCGCAGCGCGAGGTGAGAGCGCGCTGAAGGGGAGAACGAGCTCGGACGCTGGACGAC 1053
QY 341 PheHisAspLeuLeuPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
Db 1054 GTGCGAGGAGCTTAAGTACATGAATTCGGTGGTGAAGGAGACCATAGGATGACCCCTCCG 1113
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
Db 1114 ATCCCGTTGATCCCGAGATCATGCAGAGAAGAATCGAGGTCAACGGGTACACGATTCCG 1173
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db 1174 AATAAGGCCAGAAATCATGATCAACGTGTGTCATGGGTAGGAATCTCTCTACTGGGAA 1233
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
Db 1234 AAACCCGAGACCTTTTGGCCCGAAGGTTTGACCAAGTCTCAGGAGATTTCATGGGAAC 1293
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
Db 1294 GATTTTCGAGTTTCCATCCATTTGGAGCTGGAAGAAGAATCTGCCCGTTTGAATTCGGG 1353
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1354 TTGGCAAAATGTTGAGGTCCCATTTGGCAGAGCTTCTTTACCACTTCGACTGGAAGTTGGG 1413
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1414 GAAGGAATGAACCTTCCGATATGGACATGTCTGAGGCAAGAGGCTTACCGGAATAAGA 1473
QY 481 LysIleAspLeuLysLeuIleProIleProTyrGlnValSer 494
Db 1474 AAGACAATCTTCTACTCGTTCCACACACCTACGATCCTTCC 1515

;; GENERAL INFORMATION:
;; APPLICANT: Croteau, Rodney B.
;; APPLICANT: Lupien, Shari L.
;; APPLICANT: Karp, Frank
;; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
;; THE PRODUCTION OF LIMONENE HYDROXYLASES
;; NUMBER OF SEQUENCES: 58
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
;; ADDRESS: PLIC
;; STREET: 1420 Fifth Avenue, Suite 2800
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98101

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/881,784
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Shelton, Dennis K.
;; REGISTRATION NUMBER: 26,997
;; REFERENCE/DOCKET NUMBER: WSUR19777
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 224-0718
;; TELEFAX: (206) 224-0779

;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1762 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; ORIGINAL SOURCE:
;; ORGANISM: Mentha spicata
;; INDIVIDUAL ISOLATE: cDNA encoding
;; INDIVIDUAL ISOLATE:
;; IMMEDIATE SOURCE:
;; CLONE: pSM12.2

;; NAME/KEY: misc feature

;; LOCATION: 558..1212

;; OTHER INFORMATION: /product= "Probe LH-1 (Figure 4A)"

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: 39..538

;; OTHER INFORMATION: /product= "Probe LH-2 (Figure 4A)"

US-08-881-784-5

Alignment Scores:

Pred. No.: 1.03e-98 Length: 1762

Score: 983.50 Matches: 199

Percent Similarity: 62.1% Conservative: 108

Best Local Similarity: 40.3% Mismatches: 170

Query Match: 38.2% Indels: 17

DB: 3 Gaps: 7

US-10-759-813-2 (1-500) x US-08-881-784-5 (1-1762)

QY 12 LeuileSerPheLeuValLeuValLeuValValMetArgLeu----- 27

Db 32 CTTTGTGGCAATATATATCTTTGGCAACTACATCGTATCCTCTCTAATCAACCAA 91

QY 28 TrpLysLys-----GlnAsnProProGlyProTrpLysPheProLelle 43

Db 92 TGGCGAAATCGAAATCCCAACAAACCTACCTCCGAGCCCTCCGAAAGTCCGGTGATC 151

QY 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63

Db 152 GCCCACCTC---CAC---TTCTGTGGGAGGGCTTCCCGACAGCGTGTTAGGAGCATA 205

QY 64 AlaGlnIleTrpGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValle 83

Db 206 GCCCAGAAGTACGGCGCGTGGCGACGTGGGGAAGTGTACTCGTGTGTGTG 265

QY 84 SerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103

Db 266 TCGTCCGCGAGGAGCGAAGCAGCGATGAAGGTCTGGACCCCGAACTTCGCCACCGG 325

QY 104 ProIleValLeuAspAlaGlnIleValPheTrpAsnArgLysAspValPheAlaSer 123

Db 326 TTCGACGCGCATCGGTCGAGGACCATGTGTACGACAAAGATGACATCATCTTCAGCCCT 385

QY 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143

Db 386 TACAACGATCACTGGCGCCAGATCGGAGGATCTCGGTGACAGAGCTGCTGAGCCCGAAG 445

QY 144 LysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu 163

Db 446 AACGTCAGTCTCTCGGGTACATAAGGCAGGAGGATCGAGCGCTCATCGCGTGTCTC 505

QY 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTrpGlyIleIleSer 183

Db 506 GGGTCGTGGGGGAGCCCGTGCAGTGACGGAGGAGGTGCGAAAGATGTCGTGTGTC 565

QY 184 IleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeuSerValAla 203

Db 566 GTCGTGTGCAGGCGCGCTTCGGGAGTGTCTCAAGGACCAAGGT-----TCGTGTGGCG 619

QY 204 AspAlaValAsnGluAla-----AlaThrSerPheGlyThrAlaAspAlaPhePro 220

Db 620 GAGTTGGTGAAGGAGTCTCGCTCGCATTCGCGGTTTGAGCTGCGCGATCTCTACCCT 679

QY 221 ThrTrpLysLeuLeuHisTrpIleGlyAlaGluSerLysProArgLeuHisGln 240

Db 680 TCCTCATGGCTCTCAACTGCTTAGCTTGAACAAGTACAGGTTGACAGGATGCGCCGC 739

QY 241 GluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys-----Pro 258

Db 740 CGCTCGATCATCTCTTGATGGGTCTCTGAGGAGCATAGGAGAGCAAGAGCGCGCAG 799

QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuLysAsnGlyAsnValPro 278

Db 800 TTTGGAGCGAGGACATCGTCGAGCTTCTTTTCAGGATGCAGAGGGCAGCCACATCANA 859

QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298

Db 860 ATTCCCATTAATTCGAATTCGATCAAGGGTTTCATTTTCGACACCTTCTCCGCGGAGCT 919

QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318

Db 920 GAAACGCTCTCGACGACCATCTCATGGGCGTTGTTCGAACTGATGAGGAATCCGCGAAG 979

QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338

Db 980 ATGGCCAAAGTTCAGCGGAGGTAAAGAGGCGCTCAAGGGGAAAGACAGCTCGTGGATTG 1039

QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHis 358

Db 1040 AGCGAGGTGCAAGAGCTAAATACCTGAGATCGGTGTTAAAGGAGACTCTGAGGCTGCAC 1099

QY 359 ProProValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGlu 378

Db 1100 CCTCCCTTTCCATTAATCCCAAGACATCCAGGGGAAGATGCGAGGTTAACCGGTACACG 1159

QY 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398

Db 1160 ATTCGCGCAAACTAGAAATCTTCATCAACGTCTGGCTATCCGAGGGATCCCAATAC 1219

QY 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLys 418

```
Db      1220 TGGGAAGATCCCGACACCTTCGCGCCCTGAGAGATTCGATGAGGTTTCCAGGATTTTCATG 1279
QY      419 GlyThrThrPheGluLeuValProPheGlyValGlyArgLeuCysProGlyIleThr 438
Db      1280 GGAACAGATTTCAGTTTCATCCCATTCGGGGCGGTGGAAGAATCTGCCCGGTTCATAC 1339
QY      439 SerAlaIleThrAsnLeuGluTyrValIleAsnLeuLeuTyrHisPheAsnTrpGlu 458
Db      1340 TTCGGGCTGGCAAAATGTTGAGATCCCATTTGGCGCAACTGCTCTACCACTTCGACTGGAA 1399
QY      459 LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeu 478
Db      1400 TTGCCCAAGGAATGACTGATGCCGACTTGGACATGACGAGACCCAGGCTCTTCTGGG 1459
QY      479 ArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db      1460 CCAAAAAGAAAATGTTTGGTTGGTCCACACTCTATATA 1501

RESULT 12
US-09-292-768-1
; Sequence 1, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Kaip, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wbur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1762
; TYPE: DNA
; ORGANISM: Mentha spicata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1507)
US-09-292-768-1

Alignment Scores:
Pred. No.:      1,03e-98      Length:      1762
Score:          983.50      Matches:      199
Percent Similarity: 62.1%      Conservative: 108
Best Local Similarity: 40.3%      Mismatches:  170
Query Match:     38.2%      Indels:       17
DB:              3          Gaps:          7

US-10-759-813-2 (1-500) x US-09-292-768-1 (1-1762)
QY      12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeu----- 27
Db      32 CTTTTCGCGCAATTATATCTTGTGGCAACCTACATCGTATCCCTCTCTAATCAACCAA 91
QY      28 TrpLysLys-----GlnAsnProProGlyProGlyProTrpLysPheProIleIle 43
Db      92 TGGCGAAAATCGAAATCCCAACAAAACCTACCTCCGAGCCCTCCGAAGCTGCGGTTGATC 151
QY      44 GlyAsnLeuProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db      152 GGCACCTC---CAC---TTCCTGTGGGAGGGCTTCCCAGCAGCAGCTGTTTAGGAGCATA 205
QY      64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIle 83
Db      206 GCCCAGAAGTACGGGCGGTGGCGCAGTGCAGCTGGGAGAAGTGTACTCGGTGGTGCTG 265
QY      84 SerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db      266 TCGTCGGCGGAGCGACGGAAGCGGCGATGAAGGTCTGGACCCCGAACTTCGCCGACCGG 325
```

```
QY      104 ProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
Db      326 TTCAGCGCATCGGTCCAGGACCATGTGTGTCAGCAAAAGATGACATCATCTTCAGCCCT 385
QY      124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db      386 TACAACGATCACTGGCGCCAGATCGCGAGGATCTCGGTGACACAGCTGCTGAGCCCGAAG 445
QY      144 LysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeu 163
Db      446 AAGCTCAGGTCTTCCTCGGTACATAAGCGCAGGAGAGATCGAGCGCTCATCCGCTGCTC 505
QY      164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db      506 GGGTCGTGGGGGAGCGCGGTGACGTGACGAGGAGGTGTCGAAGATGTCGTGTGTGTC 565
QY      184 IleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeuSerValAla 203
Db      566 GTCGTGTGCAGCGCGCGGTTCGGGAGTGTCTCAAGGACCAGGGT-----TCGTTGGCG 619
QY      204 AspAlaValAsnGluAla-----AlaThrSerPheGlyThrAlaAspAlaPhePro 220
Db      620 GAGTTGGTGAAGGAGTCTGCTGGCATTTGGCGTCCGGGTTTCAGCTGGCGGATCTCTACCT 679
QY      221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
Db      680 TCTCATGGCTCTCTCAACTGCTTAGCTTGAACAAGTACAGGTTGACAGGATGCGCCCGC 739
QY      241 GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys-----Pro 258
Db      740 CGCCTCGATCACATCCTTCATGGTTCCTGGAGGAGCATAGGAGAGAAGAAGCGCGCAG 799
QY      259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db      800 TTTGGAGCGCAGGACATCGTCGACGTTCTTTTCAGGATCGCAAGAGGCGACATCAAA 859
QY      279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db      860 ATTCCATTACTTCCAAITGGCATCAAGGGTTTCATTTTCGACACCTTCTCCGCGGAGCT 919
QY      299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db      920 GAAACGTCTTCGACGACCATCTCATGGGCGTTGTCGGAATCTGATAGGAATCCGCGGAAG 979
QY      319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db      980 ATGGCCAAAGGTGCGCGCGAGGTAAGAGAGCGGCTCAAGGGAAAAGACAGTCGTGGATTG 1039
QY      339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db      1040 AGCGAGGTGCAAGAGCTAAATACCTGAGATCGGTGTTTAAAGAGAGACTCTGAGGCTGCGAC 1099
QY      359 ProProValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGlu 378
Db      1100 CCTCCCTTTCATTAATCCCAAGACAATCCAGGGAAGATCGAGGTTAAACGGGTACACG 1159
QY      379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db      1160 ATTCCGGCCAAAACATAGAATCTTCATCAACGTCGTGGGCTATCGGAAGGGATCCCCAATAC 1219
QY      399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
Db      1220 TGGGAAGATCCCGACACCTTCGCGCCCTGAGAGATTCGATGAGGTTTCCAGGATTTTCATG 1279
QY      419 GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db      1280 GGAACAGATTTCAGTTTCATCCCATTCGGGGCGGTGGAAGAATCTGCCCGGTTCATAC 1339
QY      439 SerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLysThrHisPheAsnTrpGlu 458
Db      1340 TTCGGGCTGGCAAAATGTTGAGATCCCATTTGGCGCAACTGCTCTACCACTTCGACTGGAA 1399
```


Db 1400 TTGCCAAGAATGACTGATGCGGACTTGGACATGACGAGACCCCGAGGTCTTCTGGG 1459
 QY 479 ArgLysIleAspLeuLysLeuIleProIleProTyrGln 492
 Db 1460 CCAAAAAGAAAAGTGGTCTGGTTCCTCCACACTCTATAAA 1501

RESULT 14

US-09-292-768-63
 ; Sequence 63, Application US/09292768
 ; Patent No. 6194185
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney B
 ; APPLICANT: Lupien, Shari L
 ; APPLICANT: Karp, Frank
 ; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
 ; FILE REFERENCE: weur13463
 ; CURRENT APPLICATION NUMBER: US/09/292,768
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 08/881,784
 ; EARLIER FILING DATE: 1997-06-24
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 63
 ; LENGTH: 1762
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: computer-generated nucleic acid sequence
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (20)..(1507)
 ; OTHER INFORMATION: Computer-generated nucleic acid sequence encoding
 ; OTHER INFORMATION: limonene-6-hydroxylase variant
 US-09-292-768-63

Alignment Scores:

Pred. No.: 3-7e-98 Length: 1762
 Score: 978.50 Matches: 199
 Percent Similarity: 61.9% Conservative: 107
 Best Local Similarity: 40.3% Mismatches: 171
 Query Match: 38.0% Indels: 17
 DB: 3 Gaps: 7

US-10-759-813-2 (1-500) x US-09-292-768-63 (1-1762)

QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValValMetArgLeu----- 27
 Db 32 CTTTTCGCGCAATTATATCTTGTGGCAACCTACATCGTATCCCTCTCTAATCAACCAA 91
 QY 28 TrpLysLys-----GlnAnProProGlyProTrpLysPheProIleIle 43
 Db 92 TGGCGAAAATCGAAATCCCAACAAAACCTACTCCGAGCCCTCCGAAGCTGCGGTGATC 151
 QY 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
 Db 152 GGCACCTC---CAC---TTCTGTGGGAGGCGTTCCTCCAGCAGCAGCTGTTAGGACATA 205
 QY 64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIle 83
 Db 206 GCCCAGAAGTACGGGCGCGTGGGCGCAGTGCAGCTGGGAGAGTGTACTCGGTGGTGTGCTG 265
 QY 84 SerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
 Db 266 TCGTCGGCGAGGACGACGAGGCGGATGAAGGTGCTGGACCCCGAACTTCGCGCAGCGG 325
 QY 104 ProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
 Db 326 TTCGACGGCATCGGTTCAGGACCATGTGTACGACAAAGATGACATCTTCAGCCCT 385
 QY 124 TyrGlyAspHisTrpArgGlnMetLysIleIleTrpIleLeuGluPheLeuSerAlaLys 143

Db 386 TACAACGATCATTGGCGCCAGATCGGAGGATCTGCGTGACAGAGTCTGTGAGCCCAAG 445
 QY 144 LysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu 163
 Db 446 AACGTCAGGTCTTTCGGGTACATAAAGCAGGAGATCGAGGCGCTCATCCGGCTGTC 505
 QY 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleIleTyrGlyIleIleSer 183
 Db 506 GGGTCGTGGGGGAGCGCGGTGCGACGTACGAGGAGGTGCGAAGATGTCGTGTGTGTC 565
 QY 184 IleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeuSerValAla 203
 Db 566 GTCGTGTGCAGGCGCGGTTCGGAGTGTCTCAAGGACCCAGGT-----TCGTTCGCG 619
 QY 204 AspAlaValAsnGluAla-----AlaThrSerPheGlyThrAlaAspAlaPhePro 220
 Db 620 GAGTGTGTAAGGAGTCTGCTGGCATTTGGCGTTTGGGTTGAGCTGGCGGATCTCTACCT 679
 QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
 Db 680 TCCTCATGGCTCTCAACCTGCTAGCTTGAACAGGTACAGTTGACAGAGATGCGCGCG 739
 QY 241 GluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaLeuLys-----Pro 258
 Db 740 CGCCTCGATCACATCTTGATGGTTCTGGAGGACATAGGAGAGAGAGCGGCGAG 799
 QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValPro 278
 Db 800 TTTGGAGGCGGAGACATCGTCGACGTCTCTTTTTCAGGATGCAGAAAGGACGACATCAA 859
 QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
 Db 860 ATTCCCTTACTTCCAAATTCATCAAGGGTTTCATTTTCGACACCTTCTCCGCGGAGCT 919
 QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
 Db 920 GAAACGTCTTCGACGACCATCTCATGGCGTTCGCGAACTGATGAGGAATCCGCGCAAG 979
 QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
 Db 980 ATGGCCCAAGTCCGCGGAGGTAAAGAGGCGCTCAAGGGGAAAGACAGTCGTGGATTTC 1039
 QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
 Db 1040 ACGGAGTCAAGAGCTAAATACCTGAGATCGTGTGTTAAGGAGACTCTGAGGCTGCAC 1099
 QY 359 ProProValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGlu 378
 Db 1100 CTTCCCTTTCCATTAATCCAAAGACAATCCAGGGAAGATGCGAGGTTAAGCGGTACACG 1159
 QY 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
 Db 1160 ATTCCGCGCAAACTAGAAATCTTCATCAACGCTGGGCTATCGGAAGGATCCCAATAC 1219
 QY 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
 Db 1220 TGGGAGATCCGACACCTTCGCGCTGAGAGATTCGATGAGTTTCCAGGGATTTCATG 1279
 QY 419 GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
 Db 1280 GGAAACGATTCGAGTTTCATCCATTCGCGGCGGTTCGAAGAATCTGCCCGGTTCATAT 1339
 QY 439 SerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGlu 458
 Db 1340 TTCGGGCTGGCAAAATGTTGAGATCCCATTTGGGCGCACTCTCTACCATTCGACTGAAA 1399
 QY 459 LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeu 478
 Db 1400 TTGCCACAAAGGAATGACTGATGCCACTTGGACATGACGAGACCCCGAGGTCTTCTGGG 1459
 QY 479 ArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492

```
Db      1460 CCAAAAAAGAAAAATGTTTGTGTTGTTCCACACTCTATAA 1501
RESULT 15
; Sequence 3, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminezky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1691 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..1545
US-08-948-564-3

Alignment Scores:
Pred. No.: 4,48e-98 Length: 1691
Score: 977.50 Matches: 207
Percent Similarity: 62.7% Conservative: 112
Best Local Similarity: 40.7% Mismatches: 165
Query Match: 38.0% Indels: 25
DB: 3 Gaps: 10

US-10-759-813-2 (1-500) x US-08-948-564-3 (1-1691)
QY      1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleGlyAsnLeuValLeuIle 20
Db      22 ATGGAGCTTCACACACACACCCCTTCTCTATTACTTACTTACCTCCATTCTCTATT 81
QY      21 LeuValValMetArgLeuTrpLysLysGlnAsnPro----- 33
Db      82 TTCTTGGTCTTCAAAATAGTTCAAGATCGGATTCCAAAACCTCTCTACCTGCMAA 141
QY      34 ---ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuThr 52
Db      142 TTGCCCCCGAGCCAGGACACTACTCTCATAGGGAACATA---CACCAGATTGTGGC 198
QY      53 SerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleIleTrpGlyProValMetSer 72
Db      199 TCACTGCCGGTTCATTACTTATAAATAATTTGGCAGATAAGTATGGTCCATTATATGCAT 258

QY      73 LeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGluAlaAlaLysGluVal 92
Db      259 CTAATAACTAGGAGAGGTGTCACACATCATAGTCACTTCCCAGAAATAGCCCAAGAGATT 318
QY      93 MetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleVal 112
Db      319 ATGAAGACACATGATCTCAACTTCTCTGATAGCCAGACTTTGTATTGTCTAGATAGTT 378
QY      113 PheTrpAsnArgLysAspValLeuPheAlaSerTrpGlyAspHisTrpArgGlnMetLys 132
Db      379 TCTTACAACGGTCTGGCATTGTCTTCAGTCAACATGAGACTATTGGAGGCACTAAGA 438
QY      133 LysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSerArgLeuLeuArg 152
Db      439 AAGATATGCACAGTAGAGTTACTAACAGCAAGCGCGTCAGTCTTTTCGGTCCATAAGA 498
QY      153 GluGluGluMetGluAspAlaIleThrPheLeuArgSerLysAla-----Gly 168
Db      499 GAAGAGAGGTGGCAGAACTAGTTAAAAAATAGCTGCACTGCAAGTGAAGAAGGGGG 558
QY      169 SerProValAsnIleThrLysIleIleTrpGlyIleIleIleSerIleMetIleArgThr 188
Db      559 TCCATTTTAAATCTCACCAGAGCATTCTCAATGACTTTTGGGATAGCGCAGCGG 618
QY      189 SerValGlyAsnCysLysGlnLysGluArgLeuSerValAlaAspAlaValAsnGlu 208
Db      619 GCTTTTGGTAAA---AAGAGCAGATACCAACAAGTGTTCATATCAACATGTCATAACAA 675
QY      209 AlaAlaThr-----SerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHis 226
Db      676 TTGATGCTTCGGGAGGGTTTCTCTGCTGATCTCTATCTCTCTCTCTAGTAGAGTTTCAA 735
QY      227 TyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIleLeu 246
Db      736 --ATGATGGGGCGCGCGGAACTTCAAAAAGTGCATAGAGTGCACAGATAGGTTGTTG 792
QY      247 GluGluIleLeuAsnGluHisLys---AlaAsnLysProPheGlu-----Ala 261
Db      793 CAAGACATCATCGAGCAGACAAAAAATAGAAACAGACGAGCGAGCGGTGAAGCAGTG 852
QY      262 AspAsnLeuMetAspValLeuAsnLeuGlnLysAsnGlyAsnValProValProVal 281
Db      853 GAAGATCTAGTTGATGTTCTCTCAAGTTTCAAAAGGAATCGAA-----TTTCGCTTG 906
QY      282 ThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThr 301
Db      907 ACTGATGACAAACATTAAAGCCGTCATCCAGGACATATTCTTGTGGAGGCGGAAACATCA 966
QY      302 SerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLys 321
Db      967 TCTTCTGTGTGGAATGGGGATGTCAGAAATTTGATAAGAAACCCGAGGGTGATGGAGAA 1026
QY      322 AlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPhe 341
Db      1027 GCACAGCAGAGGTGCGAAGAGTGTATGATACCAAGGGATATGTGGATGAGACAGAAATTG 1086
QY      342 HisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProProVal 361
Db      1087 CACCAATTTGATATACTTAAAGTCCATCATCAAGAAACCATGAGGTATACATCCACTGTG 1146
QY      362 ---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGluIleHis 380
Db      1147 CCATTGTTAGTTCCTAGAGTAAGTAGAAAGGTGCCAATCAATGGATATAGATACCC 1206
QY      381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db      1207 TCTAAGACTAGGATCATTAATGCTTGGGCAATTTGGAAGGAATCCTAAGTATTGGGT 1266
QY      401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLysGlyThr 420
Db      1267 GAAACTGAGAGTGTAAACCTGAGAGGTTTCTTAATAGCTCCATTGATTTTGGGGCACA 1326
```

```
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
Db 1327 GACTTTGAATTATCCCATTTGCTGCTGGAAGAGGATCTGCCCGGCATTACATTGCC 1386
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1387 ATACCCAAACATTGAGTTGCCACTTGCTCAGTTACTTTACCACCTTGATTGGAAGCTTCCC 1446
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1447 AATAAATGAAGAATGAAGAACTGACATGACGAGTCAAAATGGAATTACTTTACGAAGA 1506
QY 481 LysIleAspLeuLysLeuIleProIle 489
Db 1507 CAAAATGACCTCTGCTTGATTCCTATT 1533
```

Search completed: May 31, 2006, 08:41:39
Job time : 329 secs

This Page Blank (uspto)